

Towards the clinical implementation of online adaptive proton therapy

Strategies for automated and robust proton treatment planning

Elena Borderias Villarroel

A thesis presented for the degree of
Doctor of Biomedical and Pharmaceutical Sciences



Molecular Imaging, Radiotherapy and Oncology (MIRO)
Institut de Recherche Experimentale et Clinique (IREC)
Universite catholique de Louvain (UCLouvain)
Belgium



Elena Borderias Villarroel was supported by Fonds Baillet Latour

Acknowledgements

This manuscript collects most of the work I have carried out during four fantastic years of Ph.D. It has been an exciting journey, with intense moments, fascinating discussions, and some peaks of stress. Although I have invested many hours in this project, I was not alone. In the following (probably too few) paragraphs, I will express my gratitude to all those who directly or indirectly have contributed to it.

First, I would like to thank the people who made it possible and enjoyable: my promoter Edmond, my co-promoter Xavier, the head of the MIRO laboratory John, and my colleague and friend Ana. Especially to Edmond, thank you for sharing your enthusiasm for science and research, being very responsive, and providing valuable and quick feedback. I am grateful for your patience and advice when I got stuck in the writing periods and, above all, for your support and efforts in making for me, particularly during the difficult pandemic moments. You were a great promoter. Also, a big thank you to Xavier for transmitting your experienced clinical point of view so that the clinical relevance of the project would not get overshadowed by the technical details. Of course, thank you also to John for the calm you convey to the whole lab, your particular jokes, and for caring for all of us. And finally, to Ana, you have been not only a role model to me but a friend. Thank you for making me discover that there are many other things that I am not only passionate about but also good at.

All my gratitude also to the colleagues I have met in MIRO. You created in the lab a place with a friendly and happy atmosphere. I have very much enjoyed all the conversations we shared around a table full of food. I will never forget the *fonclettes*. A special thank you to Gregory for sharing this journey with me from beginning to end while making it extremely fun, and to Sara, the best companion in learning and frustration. *Un grand merci à vous.*

I also want to thank the members of my Ph.D. committee. They have followed my work during these years (Antje and Guillaume) and have provided helpful advice. Thank you also to the invited members for taking the time to read this manuscript in detail.

It was a matter of luck that I ended up doing my Ph.D. at the MIRO lab. How lucky I was. These years have brought me a unique experience that made me grow personally and professionally both in Belgium and abroad. I want to thank Vicki and Wouter for their welcome at the MAASTRO Clinic and for their efforts in carrying on the collaboration even though we had difficulties due to the arrival of the pandemic. I am also very grateful for the opportunity of spending almost three months at RaySearch Laboratories. Thank you to Rasmus, Mats, my supervisor Albin and the whole Research department for the warm welcome. The nearly three months in Stockholm were incredibly intense. They gave me the opportunity to put into practice everything I had learned, thus allowing me to finish my thesis in the best possible way.

During the last four years, and after a quite fun quarantine period I could (re)discover wonderful people around me. To my strong core, thank you for the coziest and craziest moments and above all for the hugs that have made Brussels a second home for me. *Ευχαριστώ, Thank you, Gràcies.*

Por último, agradezco infinitamente el apoyo de mis amigas de siempre, os he sentido muy cerca aunque estuvierais lejos. También dar las gracias a las amistades forjadas en mis ciudades de acogida: Argelès-Gazost, Zaragoza y Rennes, todos habéis contribuido a poner las piedras de este camino. No me olvido tampoco de aquellos que han confiado en mí y con los que he compartido momentos de pura felicidad. Para acabar, gracias a mi familia, por el apoyo en los buenos y malos momentos y sobre todo por ofrecerme siempre un lugar seguro al que volver.

Abstract

The goal of radiation therapy is to deliver a sufficient dose to cancer cells while keeping the exposure of healthy tissue as low as possible to avoid toxicity. Conventional radiotherapy has benefited from technological advances to deliver more conformal doses to the tumor, yet these techniques are still constrained to the natural interaction of photons with matter. Unlike photons, protons slow down as they pass through the patient, releasing most of their energy at the end of their trajectory in the form of what is known as the Bragg peak. By taking advantage of this feature, it is possible to precisely localize the dose peaks in the tumor, reducing the exposure of surrounding healthy tissue. Radiation therapy treatments are planned based on a pre-treatment image to simulate the patient's anatomy. However, patients are not static, and variations in the anatomy, such as weight loss, are frequent during a treatment course of several weeks. Proton therapy is especially sensitive to anatomical changes along the proton beam path, which alter the position of the Bragg peaks and can strongly deteriorate the initially planned dose. The adaptation of the initial plan is therefore required to preserve the quality of the treatment. Online adaptation in proton therapy, with the patient in the treatment position, aims to provide a rapid response to anatomical changes to compensate for proton dose distortions ensuring accurate treatment delivery. In this thesis, we investigate different strategies to automatically perform online adaptation to facilitate its clinical implementation in proton therapy. Eventually, online adaptation would allow clinicians to achieve greater control over the delivery of conformal proton dose distributions, reducing toxicity and improving patients' quality of life.

PhD committee

Promoters

Prof. Edmond Sterpin, PhD, Ir

Molecular Imaging, Radiotherapy and Oncology, Université catholique de Louvain
Laboratory of Experimental Radiotherapy, Katholieke Universiteit Leuven

Prof. Xavier Geets, PhD, MD

Radiation Oncology, Cliniques Universitaires Saint Luc
Molecular Imaging, Radiotherapy and Oncology, Université catholique de Louvain

President of the jury

Prof. Jean-Pascal Michiels PhD, MD

Radiology, Cliniques Universitaires Saint Luc, Université catholique de Louvain

Members

Prof. John A. Lee, PhD, Ir

Molecular Imaging, Radiotherapy and Oncology, Université catholique de Louvain
École polytechnique de Louvain, Université catholique de Louvain

Dr. Antje Knopf, PhD

University of Applied Sciences and Arts Northwestern Switzerland

Dr. Guillaume Janssens, PhD, Ir

Ion Beam Applications S.A. (IBA), Louvain La-Neuve, Belgium

Invited members

Prof. Wouter Van Elmpt, PhD

Department of Radiation Oncology (MAASTRO)

Maastricht University Medical Centre, Maastricht, The Netherlands

Prof. Maarten Lambrecht PhD, MD

Department of Radiotherapy Oncology,

UZ Gasthuisberg, Leuven, Belgium

List of publications

Peer reviewed publications

Online adaptive dose restoration in intensity modulated proton therapy of lung cancer to account for inter-fractional density changes

E. Borderías-Villarroel, X. Geets and E. Sterpin. [Physics and Imaging in Radiation Oncology](#), 15:30-37, 2020

Evaluation of the clinical value of automatic online dose restoration for adaptive proton therapy of head and neck cancer

E. Borderías-Villarroel, V. Taasti, W. Van Elmpt, S. Teruel-Rivas, X. Geets, and E. Sterpin. [Radiotherapy and Oncology](#), 170:190–197, 2022

Machine learning-based automatic proton therapy planning: impact of post-processing and dose-mimicking in plan robustness

E. Borderías-Villarroel, M. Huet, A. M. Barragán Montero, R. Helander, M. Holmstrom, X. Geets, and E. Sterpin. *Submitted to Medical Physics, October 2022*

Dose mimicking based strategies for online adaptive IMPT in head and neck cancer

E. Borderías-Villarroel, A. Fredriksson, X. Geets and E. Sterpin. *Manuscript submitted to Physics Medicine and Biology (PMB), December 2022*

Improved healthy tissue sparing in proton therapy of lung tumors using statistically sound robust optimization and evaluation

V. Badiu, K. Souris, G. Buti, E. Borderías-Villarroel, M. Lambrecht and E. Sterpin. [Physica Medica](#), 96:62-69, 2022

International conference communications

E. Borderías-Villarroel, K. Bernatowicz, K. Souris, X. Geets and E. Sterpin, [PO-GePV-T-96](#) *Online-adaptive proton therapy: assessing accuracy of robust dose restoration in lung patients*, Annual Meeting of the European Society for Radiotherapy and Oncology (ESTRO), 2019, Milan (Italy)

E. Borderías-Villarroel, K. Bernatowicz, X. Geets and E. Sterpin, [PO-150057](#) *Online-adaptive proton therapy : robust dose restoration in lung cancer patients*, 61st Annual Meeting AAPM, 2019, San Antonio - Texas (USA)

E. Borderías-Villarroel, X. Geets and E. Sterpin, [OC-0705](#) *Dose restoration: an online adaptive strategy to fight against range uncertainties in proton therapy*, Annual Meeting ESTRO, 2020, Vienna - Austria (Virtual)

E. Borderías-Villarroel, X. Geets and E. Sterpin, [BReP-SNAP-T-146](#) *Dose restoration in lung cancer proton therapy: an online adaptive strategy to account for inter-fractional density changes*, Blue Ribbon ePoster, 62nd Annual Meeting AAPM, 2020, Virtual

E. Borderías-Villarroel, X. Geets and E. Sterpin, [PO-55646](#) *Performance of Dose Restoration in Head and Neck Cancer for Adaptive IMPT*, 63rd Annual Meeting AAPM, 2021, Virtual

E. Borderías-Villarroel, X. Geets and E. Sterpin, [PO-1549](#) *The value of online adaptive proton therapy based on robust dose restoration for head and neck*, Annual Meeting ESTRO, 2021, Madrid (Spain)

E. Borderías-Villarroel, X. Geets and E. Sterpin, [OC-0451](#) *Comparative of different dose prediction and robust mimicking strategies for automatic IMPT planning*, Annual Meeting ESTRO, 2022, Copenhagen (Denmark)

List of abbreviations

4D-CT	4 Dimensional Computed Tomography
AI	Artificial Intelligence
ART	Adaptive Radiation Therapy
APT	Adaptive Proton Therapy
BAP	Best Adapted Plan
CBCT	Cone Beam Computed Tomography
corrCBCT	Corrected Cone Beam Computed Tomography
CT	Computed Tomography
CTV	Clinical Target Volume
DAPT	Daily Adaptive Proton Therapy
DIR	Deformable Image Registration
DIBH	Deep Inspiration Breath Hold
DM	Dose Mimicking
Dp	Prescribed Dose
DR	Dose restoration
DSC	Dice Similarity coefficient
DVH	Dose Volume Histogram
GPU	Graphics Processing Unit
GTV	Gross Tumour Volume
HNC	Head and Neck Cancer
IBO	Isodose Based Optimization
IMPT	Intensity Modulated Proton Therapy
IMRT	Intensity Modulated RadioTherapy
ITV	Internal Target Volume
KBP	Knowledge Based Planning
MANIV	Mechanical Assisted Non Invasive Ventilation

MC	Monte Carlo
ML	Machine Learning
MRI	Magnetic Resonance Image
MR-Linac	Magnetic Resonance Linear Accelerator
MU	Monitor Unit
NTCP	Normal Tissue Complication Probability
OAR	Organ(s) At Risk
OADEF	Online Adaptive based on Deformable Image Registration
OADR	Online Adaptive Dose Restoration
OAML	Online Adaptive using Machine Learning dose prediction
OAPT	Online Adaptive Proton Therapy
PBS	Pencil Beam Scanning
PCM	Pharyngeal Constrictor Muscle
pCT	Planning Computed Tomography
PET	Positron Emission Tomography
PT	Proton Therapy
PTV	Planning Target Volume
QA	Quality Assurance
rCT	Repeated Computed Tomography
RIR	Rigid Image Registration
RSM	RayStation Mimicking
RTT	Radiation Therapy Technologists
sCT	Synthetic Computed Tomography
SMG	Submandibular Gland
TCP	Tumor Control Probability
TLP	Traffic Light Protocol
TPS	Treatment Planning System
VMAT	Volumetric Arc Therapy

Contents

1	Introduction	1
1.1	Context and motivation of the thesis	2
1.2	Overview of radiation therapy	6
1.2.1	A ballistic problem: TCP against NTCP	7
1.2.2	Fractionation	8
1.3	Proton therapy	9
1.3.1	Rationale for proton therapy	9
1.3.2	Treatment delivery	10
1.3.3	Sensitivity to treatment uncertainties	14
1.3.4	Margins, robust optimization and adaptation	16
1.4	Proton treatment planning workflow	18
1.4.1	Pre-treatment imaging	18
1.4.2	Image registration	20
1.4.3	Contour delineation	20
1.4.4	Treatment planning in proton therapy	21
1.4.5	Plan evaluation	25
1.4.6	Quality Assurance	28
1.4.7	Treatment delivery	28
1.5	Adaptive radiation therapy	29
1.5.1	Anatomical changes	29
1.5.2	The rationale for online adaptive proton therapy	30
1.5.3	Adaptive radiation therapy workflow	34
1.6	The role of artificial intelligence in online adaptation	38
1.6.1	A brief introduction to neural networks	38
1.6.2	Synthetic CT generation	41
1.6.3	Automatic segmentation	41
1.6.4	Automatic planning	42
1.7	State of the art in online adaptation	43

1.7.1	Online Adaptive Radiation Therapy (ART)	43
1.7.2	Online Adaptive Proton Therapy (APT)	45
1.8	Objectives and contributions	47
2	Online dose restoration against density changes in lung IMPT	51
2.1	Introduction	53
2.2	Materials and methods	54
2.2.1	Lung patients data base	54
2.2.2	Initial treatment planning strategy	54
2.2.3	Our proposal: the dose restoration (DR) workflow	55
2.2.4	The isodose based optimized plan	55
2.2.5	DR on the treatment day: image registration, contour mapping and re-optimization	56
2.2.6	Evaluation metrics	56
2.3	Results	58
2.3.1	DR improves target coverage and homogeneity index	58
2.3.2	DR increased also robustness levels	59
2.3.3	Local dose differences were reduced using DR	61
2.4	Discussion	62
3	Clinical value of online dose restoration for head and neck cancer	67
3.1	Introduction	69
3.2	Materials and methods	70
3.2.1	Patient data and dose prescription	70
3.2.2	Treatment planning strategies	70
3.2.3	Method 1: Robust treatment planning	72
3.2.4	Method 2: Dose restoration (DR)	72
3.2.5	Method 3: Full offline adaptation	73
3.2.6	Robust evaluation of the methods	73
3.2.7	Evaluation metrics	73
3.3	Results	74
3.3.1	Analysis of DVH-metrics	74
3.3.2	DR can compensate for dose distortions	75
3.3.3	Limitations of DR for non-rigid deformations	75
3.4	Discussion	78
4	Robustness in automatic IMPT planning: impact of post-processing and dose-mimicking	83
4.1	Introduction	85
4.2	Material and methods	87
4.2.1	Patient cohort and planning strategy	87

4.2.2	Dose prediction model	88
4.2.3	Post-processing of the prediction	89
4.2.4	Dose mimicking strategies	89
4.2.5	Automatic planning strategies	91
4.2.6	Evaluation metrics	91
4.3	Results	92
4.3.1	Nominal target volume coverage	92
4.3.2	Worst-case target volume coverage	92
4.3.3	Nominal dose to organs at risk	94
4.4	Discussion	97
5	Dose mimicking for online adaptation of head and neck cancer	101
5.1	Introduction	103
5.2	Material and methods	104
5.2.1	Patient cohort and treatment planning strategy	104
5.2.2	Anatomy of the day and CBCT correction	105
5.2.3	Adaptive strategies	107
5.2.4	Online adaptation strategies	107
5.2.5	Adaptation schedule	108
5.2.6	Evaluation strategies	109
5.2.7	Statistical significance	111
5.3	Results	111
5.3.1	Target coverage in corrected CBCTs	111
5.3.2	Target coverage in accumulated doses	111
5.3.3	Target coverage for re-delineated contours	114
5.3.4	OAR sparing and toxicity on accumulated doses	115
5.3.5	Analysis of adaptation rates	117
5.4	Discussion	118
6	Conclusion and outlook	123
6.1	Potentials and limitations of the proposed online APT methods	125
6.2	Application of online APT strategies	127
6.3	Towards the clinical implementation	130
6.4	Future innovations for online adaptive IMPT	136
A	Supplementary material Chapter 2	139
B	Supplementary material Chapter 3	143
C	Supplementary material Chapter 4	147
	References	157

Chapter 1

Introduction

The present chapter (**Chapter 1**) introduces the general context, motivation, and contributions from the author, followed by a description of the state-of-the-art in proton therapy and adaptive radiation therapy. This chapter helps the readers to have a general overview of the research field before losing themselves in the following chapters, where the author's contributions are presented in detail.

1.1 Context and motivation of the thesis

The last study of the GLOBOCAN indicated that there were 19.3 million new cancer cases and almost 10 million deaths in 2020. GLOBOCAN also projects an increase of 47% of cancer cases worldwide for 2040 and states cancer as an important cause of morbidity and mortality in every world region irrespectively to the level of human development [1]. However, it is not all bad news; when correction for population growth and aging is applied to the cancer data, the resulting age-standardized death rate has globally fallen by 15% from 1990 to 2019 [2]. It seems that the world is slowly progressing in the fight against cancer. This data is consistent with well-known trends like the smoking rate declining across the world and the 5-year survival rates that increased due to the earlier detection and the improvement of treatment techniques [3, 4].

A large number of resources are devoted to the development of new technologies in cancer treatment [5, 6]. Currently, the most widely used are surgery, radiation therapy, and chemotherapy, and to a lesser extent also, immunotherapy or new targeted treatments. All these possible treatments can be used as the only treatment or in combination with others. Surgery and most radiation therapy techniques are local treatments, while drug treatments (such as chemotherapy and immunotherapy) are systemic treatments that can affect the entire body. Therefore, radiation therapy has the advantage of being local (i.e., only targeting a specific area of the body) and less invasive since it does not require surgery.

Although it is difficult to give a concrete number due to the large differences between countries in Europe, around 50% of all patients diagnosed with cancer benefit from radiotherapy, being this percentage sub-optimal in most of the cases [7]. Moreover, the number of needed radiotherapy treatment courses is expected to keep increasing in the next years [8, 9].

Cancer cells grow and divide at an abnormally rapid rate due to their unregulated proliferation. Radiation therapy uses high-energy X-rays or particles to damage the cancer cell's DNA. These ionizing radiations produce various DNA lesions responsible for cell death. However, the radiation must reach the tumor cells affecting the healthy tissue along its path. The superior growth and division rate of cancer cells causes their reparation mechanisms to fail. Thus, cancer cells are more easily destroyed by radiation, while healthy cells are more capable to repair themselves and survive the treatment. Therefore, the main goal in radiation therapy is to reduce the exposure of healthy tissue as much as possible while keeping a dose to the tumor sufficiently high to ensure the eradication of tumor cells. In a nutshell, radiation therapy is always seeking the compromise between

the highest dose to the tumor and the lowest dose to healthy tissues.

The total dose of radiation can be divided into smaller doses over several days to reduce toxicity. A conventional fractionation scheme for a prescribed dose of 60 Gy is divided into 30 fractions (days) of 2 Gy each [10]. Fractionation allows the desired dose to be administered to the tumor while protecting healthy tissue. The tendency to hypo-fractionation (higher doses in fewer fractions) is rising since long treatment courses are costly and can represent a burden for the patients. Moreover, higher doses could be more efficient against radio-resistant tumors. Another way of enlarging the benefit of radiation therapy is to improve the technology to sculpt the dose to the tumor with a more conformal shape [11]. The more we can avoid the healthy tissue, the less toxicity and the more room for dose escalation, allowing for higher tumor control probability (TCP).

A large number of patients are treated with photons as ionizing radiation, which is known as conventional radiotherapy. This technology has advanced considerably from the mid-20th century to the present day achieving high-quality treatments with improved dose conformity and healthy tissue sparing, such as intensity-modulated radiation therapy (IMRT), volumetric-modulated arc therapy (VMAT) or Stereotactic Body Radiation Therapy (SBRT). However, the photons' maximum dose is located close to the entrance of the body and followed by a smooth exponential decrease. Therefore, photons can damage healthy tissue before and after they reach the tumor. On the contrary, charged particles like protons deliver the maximum dose at the end of their path (i.e., where they stop), following the Bragg peak shape (Figure 1.1).

This physical property associated with protons (i.e., the Bragg Peak) entangles a reduced dose at the entrance and almost no dose after the tumor. Dosimetric advantages, toxicity reduction, and lower radiation therapy side effects can be achieved using proton therapy [13, 14]. However, this major advantage is also one of the biggest challenges. For a specific proton energy, the Bragg Peak position will depend on the density of the traversed matter. Therefore, density changes in the tissues before the target can substantially alter the Bragg peak location (where the proton stops and deliver the maximum amount of energy), compromising the planned dose [15, 16, 17] (see Figure 1.2). This uncertainty in the range of the proton beam, also known as range error, is a relevant issue that needs to be tackled in proton therapy. Although there are other sources of treatment uncertainties that can induce range errors [18] (i.e. patient positioning, dose calculation approximations, organ motion etc.) anatomical changes have a major impact.

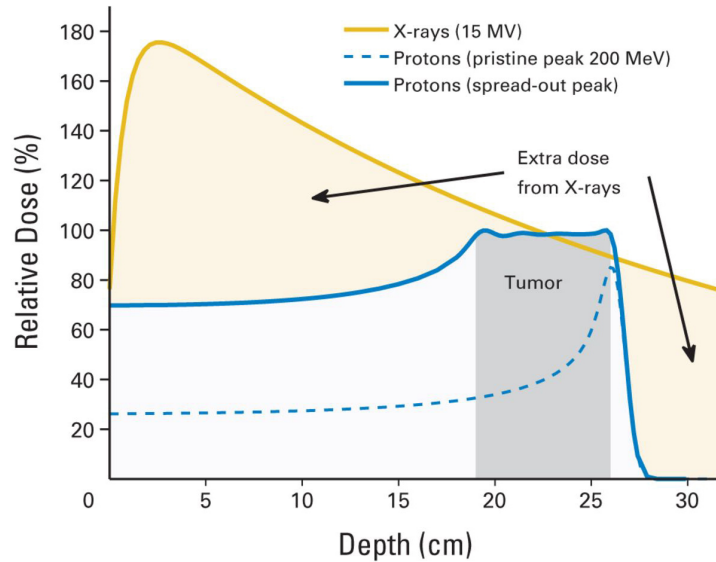


Figure 1.1: Dose deposition for conventional radiotherapy with photons versus proton therapy. Photon dose deposition in depth (yellow line), a mono-energetic proton beam dose deposition (dashed blue line), spread-out Bragg-Peak (SOBP) built from different proton energies to cover the tumor in depth (blue line). [12].

Anatomical changes have to be addressed in both conventional radiation therapy and proton therapy [20]. This is mainly due to the generation of the radiation therapy plan based on a patient image which is acquired several days before the treatment starts. Patients are not static; therefore, the optimal dose for a specific anatomy could be sub-optimal when delivered in a different anatomy a few weeks, days, or hours later. One way to ensure an accurate radiation therapy delivery (i.e., adequate dose to the tumor and no extra dose to the healthy tissue) is adapting the initial plan during the treatment course. In adaptive radiation therapy (ART), repeated patient images are regularly acquired to monitor anatomical changes and update the treatment plan accordingly. Although anatomical changes are not a proton-specific problem, proton dose distributions are much more sensitive to such variations due to the additional range errors. In the domain of photon radiotherapy, dose stability can be assumed when density changes like tumor shrinkage or weight loss occur. However, it is not always the case in proton therapy [21, 15, 16]. This is often translated into an increased need for treatment plan adaptation, especially in heterogeneous locations where anatomical changes are likely to happen, like in lung and chest [22, 23, 24] or in head and neck regions [25, 26].

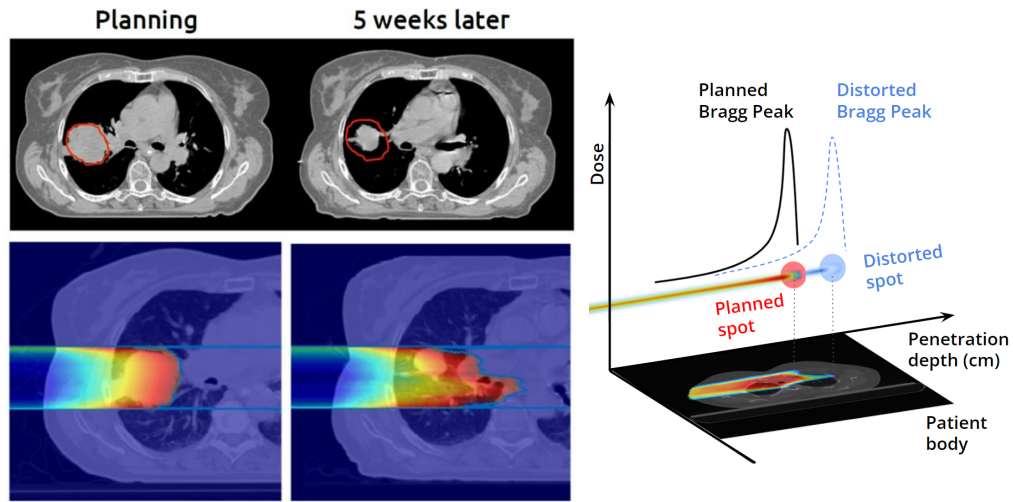


Figure 1.2: On the left, dose distortions in the presence of tumor shrinkage after several weeks. On the right, scheme of how the distortion is caused by the miss-placement of the planned Bragg Peaks due to the density changes in the proton path. Adapted from [19]

There are two main adaptation branches based on the involved timescales: on-line and offline. For online adaptation, from the moment a repeated image of the patient is acquired until the treatment is delivered, the patient remains on the treatment table at the treatment position. On the contrary, in a typical offline adaptation, a single re-planning intervention can take a few days in the clinical setting.

Offline adaptation has been available for a long time now. It has demonstrated clinical benefits and outcomes when applied in photon therapy [27, 28, 29], and it is a clear necessity in proton therapy centers [26, 30, 22]. However, not all the centers have the resources in time and human workers to implement cumbersome offline adaptation workflows with the frequency required for proton therapy. Additionally, anatomical changes, such as nasal cavity filling or air pockets, occur quite often in some locations (e.g., head and neck and pelvic regions) and can only be fully accounted for with online adaptive proton therapy (APT). Commercial solutions such as MR-Linacs and the Ethos system have become available for online ART and are generating clinical evidence about their potential benefits [31, 32]. The clinical implementation of online APT could compensate for inter-fractional anatomical changes by reducing (or even bypassing) the tedious offline adaptation workflows. However, not equivalent tools to the ones in photon therapy are available today for clinical use in proton therapy. The complexity of

the labor, the time-intensive workflow, and the lack of automation are some of the reasons hampering the widespread deployment of online APT in clinics. The main objective of this thesis is to contribute to the pathway towards the clinical implementation of online APT by proposing and evaluating fully automated proton planning workflows that could serve this purpose.

Following the idea of achieving a stable dose distribution also in proton therapy despite of density changes during treatment, online APT strategies based on the restoration of the dose have been proposed [33, 34]. The idea is to keep the planned dose distribution stable in every fraction by generating new (restored) adapted plans in a fully automated way and with minimal physician' involvement. The hypothesis is that dose restoration could already solve an important number of adaptations. The more we reduce the need for offline adaptation, the better for the patient (no delays in the treatment), and also for the service, less need of human resources. Therefore, dose restoration solutions will be studied in this thesis as a possible way to clinically implement online APT. However, although dose restoration could potentially reduce the need for offline adaptation to numbers more typical of conventional radiotherapy, it has the intrinsic constraint of not resolving non-rigid anatomical changes that would also require adaptation in photon treatments. The limitations of the method and the search for alternative solutions are also relevant topics of this work.

With the arrival of artificial intelligence (AI) in the radiation therapy field, fast and automated solutions have been proposed for automatic treatment planning. The generation of a completely new plan assisted by AI could make the planning workflow compatible with an online adaptation setting. Automated photon treatment planning using machine learning (ML) dose prediction followed by a plan optimization, known as dose-mimicking [35, 36], is one of the described pipelines in the literature. The potential of such tools for automatic proton planning to enable online APT will be also addressed in this thesis.

1.2 Overview of radiation therapy

Radiation therapy aims at eradicating cancerous cells causing DNA damage to induce cell death. The delivery of radiation therapy can be performed with different techniques. If the source is located inside the patient directly in contact with the tumor - internal radiotherapy - we refer to it as brachytherapy. Radioactive seeds can be implanted in the patient or introduced and removed (after and during treatment delivery) using catheters. In the case of nuclear medicine therapy, radiopharmaceuticals can also be injected into the patient. These drugs have the ability to migrate and get attached to the tumor to locally deposit the ionizing

radiation they emit. In the case of external radiotherapy, which is by far the dominant modality [37], tumor cells are irradiated with ionizing radiation generated by a particle accelerator. Conventional radiotherapy treatments use high-energy photon beams because of their simplicity and cost-efficiency. However, the interest in other particles, such as proton or carbon ions [38], has increased in the last years due to their particular interaction with matter, which can provide dosimetric advantages (e.g., a more conformal dose to the target and OAR sparing) [39, 40]. These benefits are crucial in external radiation therapy since its main objective is to find for every patient the best compromise between a sufficient dose delivered to the target and a low dose deposited in the healthy tissue. This is also known in radiation therapy as the ballistics problem.

1.2.1 A ballistic problem: TCP against NTCP

The tumor control probability (TCP) is defined as the probability that all cancerous cells will be killed for a given dose deposited in the tumor. Similarly, the normal tissue complication probability (NTCP) is defined as the probability that a given dose to a specific OAR will lead to complications due to the toxicity of radiation.

A TCP curve is created by plotting some measure of tumor control probability against the total dose. A TCP curve is characterized by a sigmoid shape, in which there is little dose response at very low or very high doses and a steep dose response at intermediate doses. An NTCP curve is constructed the same way as a TCP curve by plotting the probability of normal tissue injury against the dose. The therapeutic window is a figurative space between treatment failure and toxicity. The larger the window, the more likely the treatment is to be safe and effective. In order to optimize the outcome of the patient treatment, the delivered dose should therefore maximize the TCP, while keeping the NTCP as low as possible. However, due to the proximity of TCP and NTCP curves, this is not easily achievable in practice.

Several mechanisms can be implemented to modify the TCP and NTCP curves, such as using radio-sensitive and radio-protector drugs [41, 42]. It has also been shown that delivering the treatment using very high dose rates also has a radio-protector effect [43], known as the FLASH effect. Finally, the treatment delivery in several fractions [44] and the development of more conformal treatment delivery techniques to spare the maximum amount of healthy tissue (e.g., proton therapy) are also parameters the clinicians use to improve the therapeutic window.

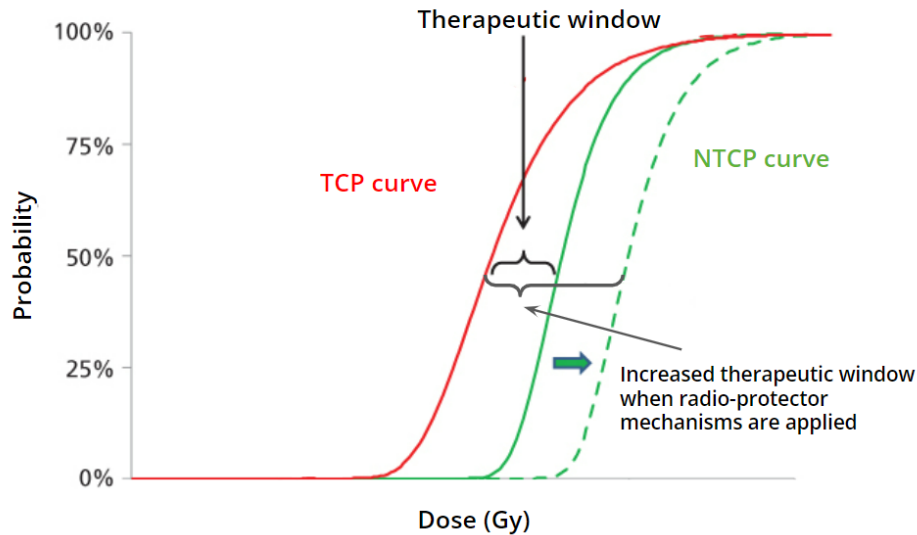


Figure 1.3: The tumor control probability (in red) and normal tissue complication probability (in green) are represented by sigmoid functions. The optimal dose should maximize the TCP and minimize the NTCP. Adapted from [41]

1.2.2 Fractionation

Fractionation in the context of radiotherapy is the process of dividing a dose of radiation into multiple “fractions”. This practice seeks to maximize the destruction of cancer cells while minimizing damage to healthy tissues. The main mechanisms which represent the rationale behind fractionation are outlined below.

- **Redistribution:** The radio-sensitivity of cells depends on their stage in the cell cycle. At the moment of treatment delivery, cancer cells are at different points in the cycle. Therefore, delivering the entire dose in a single fraction would be ineffective against a portion of the tumor cells. Dividing the total dose of radiation into multiple fractions maximizes the probability of irradiating cells when they are in the most radiosensitive period of their cell cycle.
- **Re-oxygenation:** When tumor cells are hypoxic, they are less susceptible to the indirect effects of radiation. Fractionated radiotherapy allows cells that are closer to sources of oxygen to be killed first. During the intermediate time between fractions, relatively hypoxic cells have the opportunity to improve their oxygen supply. These cells will then be more sensitive to subsequent doses of radiation in the following fractions.

- **Repair:** Normal cells have an increased ability to repair DNA damage compared to tumor cells. Therefore, splitting the total radiation dose gives healthy cells the opportunity to repair this sublethal damage between fractions. Meanwhile, cancerous cells with impaired DNA repair pathways are less able to recover from radiation damage to their DNA.

The fractionation helps to achieve the primary goal of radiation therapy, which is to deliver a homogeneous dose to the target volume while minimizing, as much as possible, the dose deposited in the healthy tissue. Nevertheless, treatment schedules are costly and can be inconvenient for patients. Alternatively, hypofractionated radiotherapy is delivered in fewer but higher dose fractions. The unique physical interaction of protons (i.e., the Bragg Peak) brings the possibility to more precisely target the tumor and allow for escalated dose delivery in hypofractionated schemes [11, 45]. Nevertheless, the combination of hypofractionation and proton therapy is quite challenging due to the increased dosimetric impact of possible anatomic variabilities for two main reasons. First, inter-fractional anatomical changes could result in distorted proton doses, which would be even more critical with increased doses per fraction. Second, since fewer fractions are delivered, the delivery errors have fewer possibilities to be averaged out in other treatment fractions. Therefore, hypofractionated schemes can potentially benefit from adaptive approaches to ensure a safe treatment delivery at higher doses per fraction [46], especially when combined with proton therapy.

1.3 Proton therapy

Proton therapy is far from being a new radiation therapy technique. The first patients were treated at Berkley Radiation Laboratory in 1954, but it was not until 1990 that the first clinical center dedicated to proton therapy opened in California. Currently, there are 109 operating centers worldwide according to the PTCOG statistics [47], 32 under construction and 29 in the planning stage. The growth of proton therapy is remarkable. While in 2007 the number of patients that benefit from proton therapy barely attained 50.000, in 2021 it was multiplied by more than five getting close to 280.000.

1.3.1 Rationale for proton therapy

Proton therapy can create more conformal dose distributions around the tumor, avoiding extra doses to the healthy tissue in the vicinity compared to photon therapy (Figure 1.1). In the case of photons, after the first build-up region up to the maximum energy posited, the dose falls off continuously through the patient's body due to the attenuation of the beam. On the contrary, protons deposit their

energy following the Bragg Peak shape. Protons start to lose energy (i.e., to reduce their speed), increasing the number of interactions with matter (ionisations) as they get closer to being stopped inside the patient. At this point, the biggest energy transfer (from proton to matter) occurs, delivering a peak of dose at the end of their path at the Bragg Peak position. When protons are stopped (with no energy left), no more ionization nor dose deposition occurs (Figure 1.1). Therefore, the normal tissue after the tumor receives less dose when irradiating with protons than it does with photons. Examples of the difference in proton and photon dose distributions are shown in Figure 1.4 for craniospinal and brain locations. Additionally, proton beams damage cell DNA in qualitatively different ways than photons (X-rays) do. It has been estimated that protons result in approximately 10% more biological damage per unit dose than photons [48].

Proton therapy has been demonstrated to have multiple benefits due to its particular physical interaction. It has the potential to minimize radiation exposure of healthy tissue [49, 14] and improve the patient's quality of life because of the reduced toxicity [13, 50, 51]. Moreover, proton therapy can provide a lower integral dose per treatment to the patient [52], and it could potentially reduce the risk of secondary cancers [53, 54]. Finally, the possibility of re-irradiation for patients with recurring forms of cancer is widened with the use of proton therapy [55].

1.3.2 Treatment delivery

In order to deliver the prescribed dose to the tumor, protons need to reach the tumor inside the patient. The depth attained by the proton depends on its energy and the traversed material. Protons with an energy of 70 and 230 MeV get to approximately a depth in the water of 4 cm and 33 cm respectively. Protons need to be accelerated inside a cyclotron (or synchrotron) to reach such energies. Afterwards, the accelerated protons are transported into the treatment room through the beamline until they reach the treatment head (or nozzle) with the desired energy to deliver a clinical 3D dose distribution. Passive scattering (PS) and pencil beam scanning (PBS) are the two available technologies for proton treatment delivery. PS uses patient and field-specific compensators to scatter the proton beam, while in PBS magnetic fields steer the proton beam to scan the desired area (Figure 1.5).

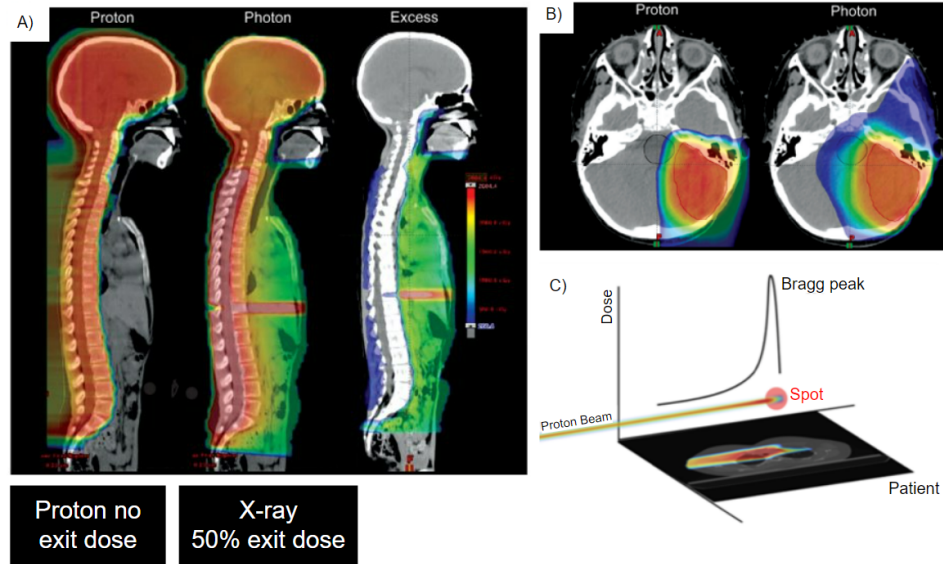


Figure 1.4: Differences in proton versus photon dose distributions for craniospinal (A) and brain (B) cases. Due to the proton dose deposition following the Bragg Peak shape (C), there is almost no exit dose when irradiating with protons compared to photons. Adapted from [56]

Although the first proton centers were equipped with PS systems, nowadays, new centers are mostly equipped with PBS systems as it is the most advanced form of treatment delivery. Contrary to PS, PBS treatment delivery does not require the use of patient-specific and field-specific compensators and collimators. This eliminates treatment delays, reduces treatment time and costs, and increases flexibility in delivery. Moreover, PBS reduces patient exposure to secondary radiation, produced when the beam hits a device, which could cause secondary malignancies [54]. PBS also enables intensity-modulated proton therapy (IMPT) delivery combining Bragg-peaks with different energies from different beam angles to achieve a more conformal dose to the tumor. IMPT has demonstrated its superiority against PS and conventional radiotherapy in terms of dose conformity, healthy tissue sparing, and degrees of freedom to shape the dose [58, 59].

PBS delivery can be especially challenging when mobile tumors are being treated. The organ motion interferes with the dynamic delivery, introducing treatment uncertainties on the Bragg Peak positions and, therefore, dose distortions with cold and hot spots. This is known as the *interplay effect*. In this thesis, interplay effects are disregarded. However, this problem has been studied extensively and several publications have demonstrated that its dosimetric effects average out over

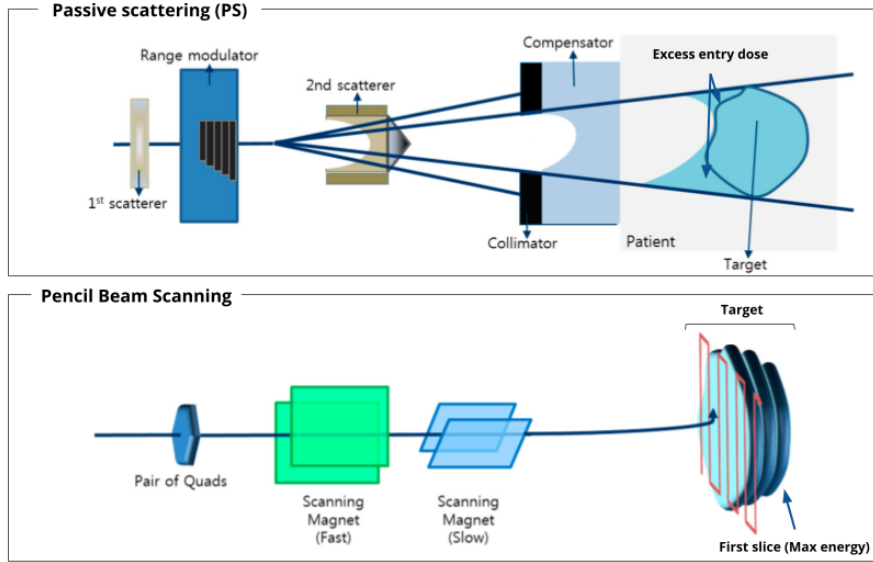


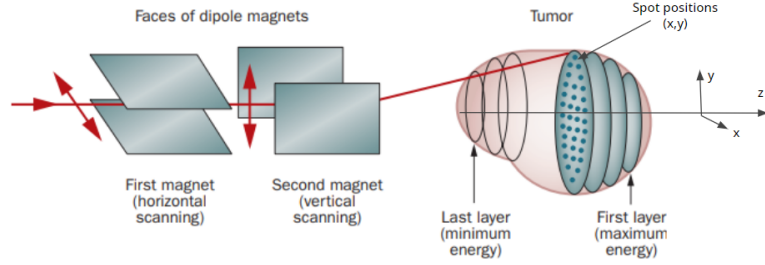
Figure 1.5: : Proton delivery techniques: passive scattering (upper figure) versus pencil beam scanning (lower figure). Adapted from [57].

the course of a conventionally fractionated treatment (i.e., 30 fractions) [60, 61], yielding a conformal overall dose distribution at the end of the treatment. For each individual fraction, interplay effects can be significantly reduced with simple techniques such as rescanning [62, 63].

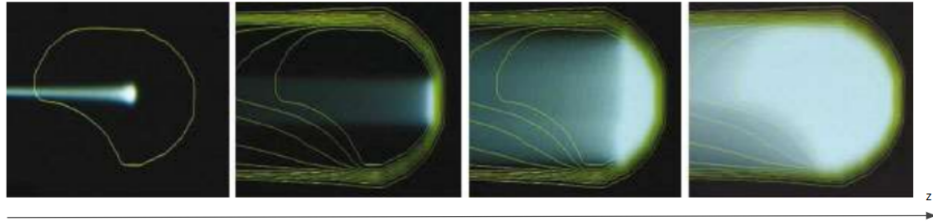
To cover the tumor in depth with PBS delivery, Bragg peaks of different energies are superposed to form the so-called Spread Out Bragg peak (SOBP) (Figure 1.1). The tumor is scanned in the beam direction by energy layers, starting with the highest energy at the most distant layer (i.e., the deepest tumor region), Figure 1.5. The desired proton energy is attained differently depending on the accelerator system. For example, in the cyclotron developed by IBA (Ion Beam Applications S.A., Louvain-La-Neuve, Belgium), the protons are always accelerated to the maximum energy (230 MeV). Their energy is degraded afterwards until the desired value. The minimum energy that can be selected is 70 MeV, which corresponds to a depth of 4 cm. If energies below this threshold want to be used to reach smaller depths (under 4 cm), a piece of material, known as the range shifter, needs to be inserted in the head of the accelerator. This device enables to decrease of the energy of the beam down to a few MeV, which is convenient for superficial tumor locations, such as tumors in the head and neck region close to the skin.

To cover the tumor laterally, each energy layer is designed as a discrete scanning grid of dose delivery positions (spots), Figure 1.6. The number of spots per layer can vary a lot, from a few spots (less than 10) to more than a hundred. The space between spots is typically around 3-5 mm, which leads to a total number of spots, including all energy layers, that can reach several thousand for regular tumor sizes.

a) PBS delivery scheme



b) Evolution of deposited dose for PBS treatment delivery



c) Example of spot map with x,y positions for the energy layer at 109 MeV in the beam eye view (BEV)

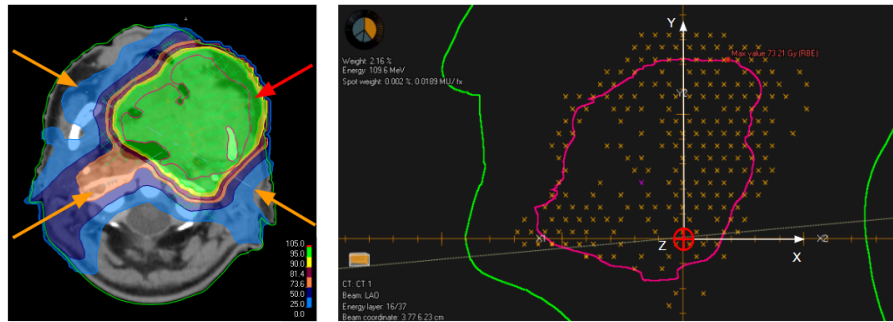


Figure 1.6: : (a) Illustration of the scanning grid energy layers and spot positions for PBS. (b) Evolution of PBS dose delivery by energy layers. (c) An example of an IMPT treatment plan in which the orange crosses correspond to the spot positions. Adapted from [64] and [65]

1.3.3 Sensitivity to treatment uncertainties

Radiation therapy treatments are planned based on a pre-treatment computed tomography (CT) image, which is supposed to serve as an estimation of the patient's anatomy. However, this CT image is often called planning-CT (pCT) and is acquired several days before the treatment starts. Thus, the pCT does not fully correspond to the real anatomy of the patient at every treatment fraction. Such estimation of the patient anatomy, together with the characteristics of proton therapy (e.g., the steep dose gradient and the Bragg Peak), makes this modality particularly sensitive to variations (i.e., uncertainties) that occur between and/or within fractions. In this section, the concept of treatment uncertainty will be introduced. A special focus will be put on proton range uncertainties which are specific to proton therapy. Other sources of treatment uncertainty will be discussed further in the manuscript when describing the proton treatment workflow in depth (see section 1.4).

In radiation therapy, uncertainties are usually divided into systematic errors (Σ) and random errors (σ) [66, 67]. Errors can be defined as any source of variability that can cause a deviation between the planned (before delivery) and the executed treatment (delivered). Systematic errors equally influence all treatment fractions. An example of systematic error is the uncertainty in the delineation of the target volume or organs at risk (OAR) in the pCT. This uncertainty is propagated to all fractions in which we assume the same patient anatomy. On the other hand, the magnitude of random errors is different for every fraction. It can highly vary during the course of the treatment depending on inter-fraction variations. Organ motion, such as breathing motion or bladder filling, are examples of random errors. Other uncertainties combine both systematic and random errors. This is the case of the patient positioning error, also known as **setup error**, which happens when trying to replicate the patient's treatment position at every fraction to ensure an accurate treatment delivery. Setup errors are a very common source of treatment uncertainty in any type of radiation therapy treatment without being specific to protons. However, in proton therapy setup errors induce proton range distortions in addition to the aiming error. This additional uncertainty related to the planned position of the Bragg peak (i.e., proton range) is specific to proton therapy. Uncertainties in the proton range occur when the proton does not stop at the estimated depth. This uncertainty is also known as **range error** and can result in distorted doses and sub-optimal treatment outcomes.

The finite depth (range) of each Bragg peak within the body is dependent on the energy of the pencil-beam and on the stopping power of the tissues the pencil-beam encounters along its path. Therefore, changes in the densities (i.e., stopping

powers) along the beam paths can alter the depths of the individual Bragg peaks and change the overall shape of the dose distribution. As a result, local over- and under-dosages can occur (see Figure 1.7). One could distinguish between inter-fractional anatomical variations, which occur from one treatment fraction to the next, and intra-fractional variations, which can happen within the same treatment fraction. Weight loss and tumor regression are gradual inter-fractional anatomical changes that can highly impact the quality of proton plans [68, 69]. Breathing motion or bladder filling are examples of much faster intra-fractional changes.

Range errors in proton treatments can also arise from dose calculation approximations and biological effects. An extensive description of range uncertainties can be found in [18]. However, these will not be directly addressed in this thesis. We will mainly focus on range errors induced by anatomical changes that can result in distortions of the proton dose distributions and potentially in sub-optimal overall treatments. The development and evaluation of strategies that can mitigate such dose distortions in IMPT plans represent the main topic of this thesis.

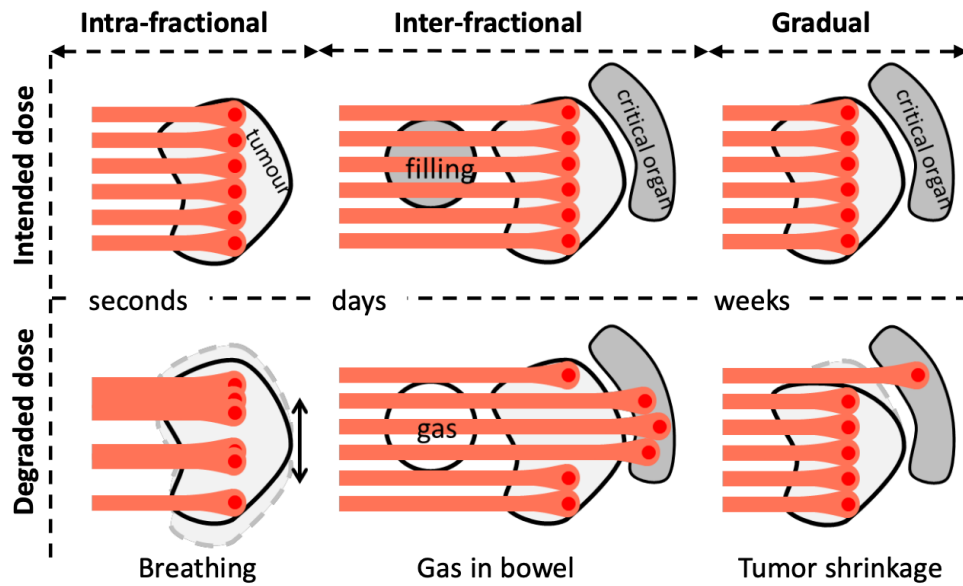


Figure 1.7: Illustration showing the effect of range distortions for different anatomical variations for IMPT delivery. Each pencil-beam is individually affected by the changes in the patient anatomy, leading to a sub-optimal dose distribution with considerable changes in the dose shape. Courtesy of M.S. Hoogeman and T. Jagt from [70]

1.3.4 Margins, robust optimization and adaptation

In order to accommodate for treatment uncertainties, such as setup and range errors, different solutions have been proposed:

- **Margins:** increasing the irradiated area around the tumor anticipates possible errors during treatment delivery. Safety margins can be modeled mathematically to account for different types of treatment errors.
- **Robust optimization:** A more sophisticated strategy is to include a limited set of pre-defined treatment uncertainties directly into the treatment planning optimization.
- **Treatment plan adaptation:** The initial dose can be actively updated in the middle of the treatment course to accommodate for patient-specific anatomical changes that occurred.

Margins

The classical way to account for setup errors and other geometric uncertainties is to add a safety margin around the clinical target volume (CTV) to build the planning target volume (PTV), following well-established margin recipes as the one proposed by Van Herk et al. [66]. However, safety margin recipes rely on the *static dose cloud approximation*, which assumes an invariant dose distribution in the presence of treatment errors. This hypothesis is widely used and approximately valid for conventional photon therapy, but it is not extensible to proton therapy [15, 16, 71] since it is invalidated by the effects of the proton range errors (Figure 1.8). This is especially true for IMPT treatments. In IMPT, the doses from different beam directions are typically heterogeneous and any variation can lead to a mismatch of the individual beam contributions, causing dramatic perturbations to the total dose distribution. Therefore, PTV-based optimization is not appropriate for IMPT. Fortunately, this limit motivated the development of more sophisticated methods, such as *robust optimization* [72, 73, 74], which takes into account the uncertainties during the optimization process.

Robust optimization

Robust optimization is designed to incorporate errors directly into the optimization algorithm. Better OAR sparing and less toxicity have been shown for robust optimization strategies compared to margins [76, 77, 78]. The objective during robust optimization is to find through an iterative process the set of plan param-

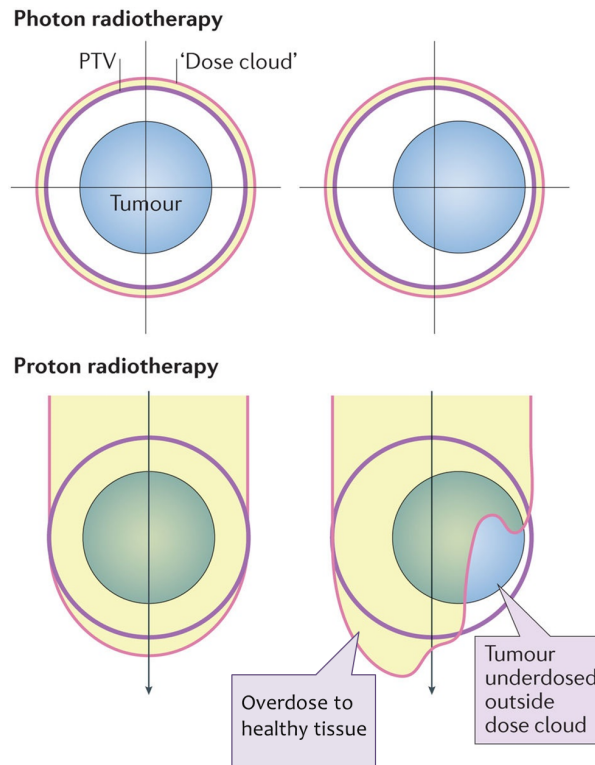


Figure 1.8: Conventional radiotherapy assumes that the dose distribution is invariant under displacements of the target (*static dose cloud approximation*), which enables obtaining enough robustness against uncertainties by simply adding a safety margin around the target (PTV). In proton therapy, this hypothesis does not hold due to the finite range of the protons, which leads to dramatic changes in the dose when the target moves. Adapted from [75].

ters that ensure that the delivered dose distribution would be minimally distorted when uncertainties are present. To do so, several uncertainty scenarios are simulated. During the optimization process, the machine parameters providing better dose distributions in all explored scenarios (i.e., less distorted) are favored, while the ones causing higher distortions are penalized. On the one hand, the more scenarios are simulated, the more robust the plan would be against errors and the lower probability of missing the tumor. On the other hand, the more scenarios are included, the more computationally expensive the process gets. Moreover, if unrealistic scenarios are explored, the resulting dose distribution unnecessarily irradiates part of the healthy tissue that could be spared with more realistic simulations. Consequently, a limited amount of uncertainty scenarios are chosen to be simulated and combined during robust optimization (e.g., range errors, setup errors, and some anatomical errors).

Adaptive radiation therapy

Both margins and robust optimization are performed at the planning stage, expecting to be robust enough during the treatment course by increasing the dose to the surrounding healthy tissue to account for uncertainties. Although both provide a certain level of robustness, it has been demonstrated that ART is still necessary to correct dose distortions induced by range errors in IMPT[26]. This is especially true for those locations which are prone to frequently experiment anatomical changes such as lung [22, 79] or head and neck [80, 25].

The presented strategies can be combined to deliver the treatment plan accurately. In current clinical practice, robust optimization is the preferred planning strategy for IMPT. When the robustness of the plan fails at compensating for the anatomical changes and the accuracy of the treatment plan is compromised, offline adaptation is applied [26, 81].

Now that proton therapy has been introduced and its main limitations explored, we will present in the next section the complete proton treatment planning workflow (section 1.4). Understanding all the steps involved in this process is necessary to comprehend the challenges of its adaptation (section 1.5), especially when performing online APT. Those familiar with proton therapy planning can skip the following section and go directly to section 1.5.

1.4 Proton treatment planning workflow

The treatment planning process is intended to generate a treatment plan that will be loaded into the treatment machine to be delivered to the patient typically in several fractions. Physicians, physicists, and all members of the radiation therapy service work together to ensure that the prescribed dose will be delivered as accurately as possible. Every step of this complex radiation therapy workflow is detailed in the following sections, specifically for IMPT. A summary of the key steps of the workflow is shown in Figure 1.9.

1.4.1 Pre-treatment imaging

The first step is to image the patient. The CT image of the patient gives access to the anatomy and electronic density information of the tissues at the moment of the image acquisition. The imaging step is key since the resulting CT image will be used for treatment planning. The anatomy of the pre-treatment CT image will not correspond to an exact representation of the anatomy of the day at every treatment fraction. However, it is the more accurate representation

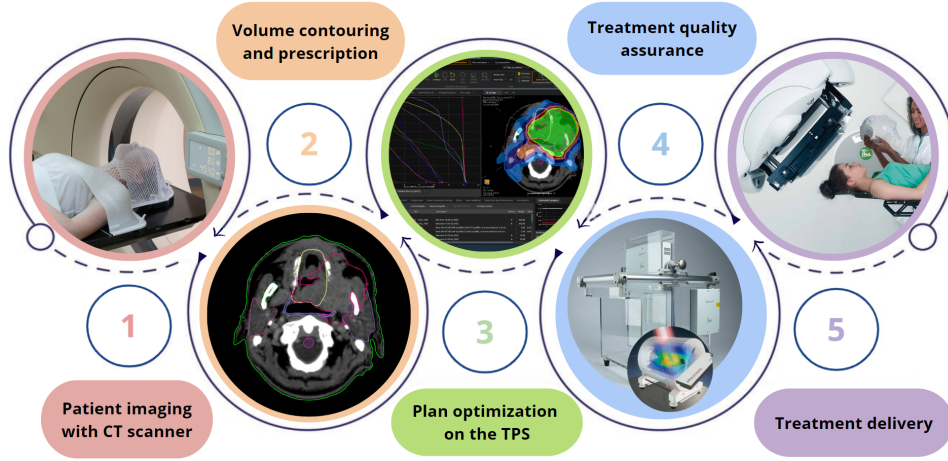


Figure 1.9: Summary of a typical proton therapy workflow from patient imaging to treatment delivery

that is available for treatment planning. Therefore a set of strategies is put in place during radiation therapy workflow to achieve the highest reproducibility and reduce the impact of uncertainties during treatment delivery.

The CT image is acquired by moving a radiation source around the patient. The source emits photons that will be differently attenuated by the distinct tissues (e.g., bone or air). By detecting the remaining photons after the patients, a 3D image can be reconstructed from the measured projections around the patient. The 3D map of the X-ray attenuation is expressed in Hounsfield Units (HU) defined as:

$$HU_{tissue} = 1000 \left(\frac{\mu_{tissue}}{\mu_{water}} - 1 \right) \quad (1.1)$$

where μ_{tissue} is the linear attenuation coefficient that describes the fraction of attenuated incident photons in a mono-energetic beam per unit thickness of a material.

The tumor area and its immediate surroundings are typically imaged. If the tumor is mobile, like lung or liver malignancies, instead of acquiring a single 3D-CT, 4D-CT can be used to characterize breathing or other organ motions. Assuming that the motion is periodic, a 4D-CT image can be reconstructed [82], providing multiple 3D-CT images at different phases of the breathing motion.

1.4.2 Image registration

Image registration algorithms are optimized to find spatial correspondence between two images or among multiple sets of images while minimizing the error (discrepancy) between them. Registration algorithms provide deformation fields that characterize the transformation of one image into the coordinate system of another one. The registration process is often used in the radiation therapy workflow to overlay two images acquired at different times and/or with different modalities. For instance, Positron Emission Tomography (PET) images are registered to CT images to combine the anatomical information encoded in the CT images with the functional information provided by the PET image. Two main types of image registration can be described depending on the operations used to obtain the deformation field from one image to the other:

- **Rigid Image Registration (RIR):** rigid transformations only include rotation and translation operations that are uniformly applied to every voxel. Rigid registration is clinically used for patient repositioning in order to align the cone beam computed tomography (CBCT) of the day (containing the anatomical information of the day) to the pCT (reference anatomy) before treatment delivery.
- **Deformable image registration (DIR):** DIR is a non-rigid image registration with a significantly greater number of degrees of freedom. Therefore, DIR can manage local distortion between two image sets (i.e., anatomical structure changes). A DIR is characterized by a deformation vector field (DVF) which describes the displacements of every voxel from one image to the other. Although DIR algorithms are not often used during the treatment planning process, they are very useful for plan evaluation during the treatment (e.g., dose accumulation) or for treatment adaptation workflows that will be introduced in section 1.5.3.

1.4.3 Contour delineation

Once the CT image is acquired, physicians delineate on it the target volumes and the OAR. Sometimes, other imaging modalities, like magnetic resonance imaging (MRI) with a better soft tissue contrast or PET images which give functional information, are used to delineate the tumor.

There are several volumes often used in radiotherapy planning. The first one is the extent of the primary tumor, known as the gross tumor volume (GTV), and corresponds to the tumor that can be seen or imaged. The second volume

is known as the clinical target volume (CTV). The CTV aims to include the not imageable microscopic infiltrations of the tumor spread or regions suspected to be attained by cancerous cells. Therefore, the CTV is usually built from the GTV as an isotropic margin of a few millimeters or corresponds to an anatomical region in which cancer cells are suspected, such as contaminated lymph nodes. In the first case, the CTV expansion is typically corrected according to the anatomic boundaries of the patient's anatomy, such as blood vessels.

To avoid cancer recurrence, the CTV must be adequately treated to ensure all cancerous cells are killed. Possible treatment uncertainties must be considered to guarantee good coverage (i.e., full irradiation) of the CTV during all fractions. As previously explained, the most common method to increase the treatment robustness against uncertainties is to expand the CTV with an additional safety margin to generate the PTV. However, in the case of IMPT planning, robust optimization based on the CTV is often preferred.

In the case of mobile tumors, a fourth target volume, known as the internal target volume (ITV), is used to account for organ motion. The ITV is constructed as the union of CTV contours delineated in the different phases of a 4D-CT image. The whole area where the tumor follows an internal movement (e.g., breathing motion) is treated by irradiating the ITV region.

CTV delineation is very challenging since it often requires clinical assessment of risk and extent of spread which is generally based on population studies instead of in a patient-specific quantification. To reduce the inter-observer variability, delineation guidelines generated by multi-national experts compile the consensus recommendations [83].

The OAR in the tumor vicinity are also delineated to optimize and evaluate the treatment plan. The more radiosensitive tissues to radio-induced side effects are mainly considered. For example, the parotids, submandibular glands, or the cochleas are important OAR for head and neck cancer (HNC), while the lungs, the esophagus, the spinal cord, and the heart are especially relevant in lung cancer.

1.4.4 Treatment planning in proton therapy

Treatment planning is based on two main pillars: dose calculation and plan optimization. For IMPT plans, the spot parameters are optimized to obtain the machine settings that can deliver the desired dose distribution. Therefore, both tasks are dependent on each other. The dose contribution of each spot and the total dose distribution are used during plan optimization to adjust the spot weights until the most optimal deliverable dose is found. In the clinical setting, treatment

planning (including dose calculation and plan optimization) and its evaluation are performed on a commercial treatment planning system (TPS).

Proton dose calculation

In order to simulate and optimize the treatment plan delivery, dose calculation in the patient is required. The total absorbed dose is defined as the total energy absorbed by a material per unit of mass. The Bethe-Bloch formula [84, 85] describes the mean energy loss per unit of distance traveled by charged particles (e.g protons), also known as the stopping power $-dE/dx$. The stopping power depends on the proton kinetic energy and the properties of the traversed material:

$$-\frac{dE}{dx} = K\rho\frac{Z}{A}\frac{z^2}{\beta^2}\left[\frac{1}{2}\ln\frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{C}{Z}\right] \quad (1.2)$$

where:

$K = 4\pi N_A r_e^2 m_e c^2$ is a constant,

N_A is the Avogadro number,

ρ is the mass density of the material,

Z is the effective atomic number of the material,

A is the effective atomic mass of the material,

z is the charge of the particle (1 for protons),

β, γ are the relativistic parameters of the particle,

I is the mean excitation potential of the material,

C is the shell correction for low energy effects,

T_{max} is the maximum kinetic energy that can be transferred to an electron:

$$T_{max} = \frac{2c^2 \beta^2 \gamma^2}{1 + 2\gamma \frac{m_e}{M} + \left(\frac{m_e}{M}\right)^2} \quad (1.3)$$

being M the mass of the particle (938.3 MeV/ c^2 for protons).

In proton therapy, the stopping power ratio (SPR) between the tissue and water is necessary for dose calculation.

$$SPR = \frac{(dE/dx)_{tissue}}{(dE/dx)_{water}} \quad (1.4)$$

Calibration curves are empirically obtained [86] to establish the relationship

between the CT image (i.e., HU values) and SPR. Such curves are often generated using the stoichiometric method proposed by Schneider et al. [87] using CT images of a phantom with several tissue-equivalent materials. The calibration curve allows the extraction of the SPR on each voxel of the patient to compute the dose. However, the conversion of a CT number to SPR has been shown to contribute to approximately 1.6% range uncertainty in proton dose calculation [88]. Dual-energy CT has been employed recently to improve stopping power predictions [89], by performing two or more CT acquisitions with different energies of the X-ray spectrum. For IMPT delivered using PBS, the two main algorithm types for dose calculation are Monte Carlo (MC) and analytical, also known as pencil beam (PB), algorithms [90].

In PB algorithms, each spot in a proton PBS beam is split into a subset of segments (beamlets) commonly referred to as pencil beams or dose kernels [91]. Multiple elementary dose kernels are measured or pre-computed in a water phantom and stored in a database. After transporting all pencil beams into the CT image, the doses are summed for each voxel resulting in a final calculated dose distribution. Classical PB algorithms use the central axis approximation. They assume that a material at a depth Z along the central axis of a pencil beam is laterally infinite. This approach makes PB algorithms less accurate in heterogeneous media, such as the lungs [18, 92, 93].

On the other hand, MC dose calculation algorithms accurately handle heterogeneous materials. MC algorithms use random number generators and probability density distributions to simulate physical interactions related to radiation transport for each particle along the beam path. With accurate physics models, the estimated dose by MC algorithms will converge to the measured dose (i.e., true delivered dose) as the number of simulated particles increases and the statistical uncertainty decreases. The drawback of MC is that it requires long computation times. That is why analytical algorithms had a quicker implementation into clinical routine, despite being less accurate. However, new fast MC codes exploit the potential of novel computing hardware, such as graphical processor units (GPUs) [94, 95] to achieve dose computation in the order of seconds or few minutes [96].

Intensity modulated proton therapy plan parameters

The treatment plan contains the machine parameters required to deliver a specific dose distribution. A specific beam model must be generated and carefully commissioned for each treatment machine and delivery technique. This process ensures that the dose planned on the TPS will correspond with the dose delivered by the machine.

The following parameters mainly characterize an IMPT plan delivered with PBS:

- **Beams:** The number and angles of the beams are manually set. An appropriate choice of both parameters can avoid critical OAR and provide more robust dose distributions. For instance, posterior beams are preferred for chest irradiation to mitigate the effect of the breathing motion. In general, an attempt is made to minimize the traversed path by protons to reduce possible range errors during the treatment.
- **Energy layers (E):** The distal part of the tumor in the beam direction (z) is covered by the highest energy layer and the subsequently lower energy layers until the more external part, i.e., closer to the skin. Energies are expressed in mega electron volts (MeV).
- **Spots positions (x, y):** Each energy layer is scanned in a discrete way stopping at specific spot positions (x, y). At this position, the dose is delivered in the form of Bragg Peak.
- **Weights (w):** The spot weight is a magnitude proportional to the number of delivered protons expressed in Monitor Units (MU). The highest the weight, the more important and the longer it takes for delivery.

Therefore, for every beam of an IMPT plan, each spot is characterized by four parameters (x, y, E, w). The beams are usually manually defined, while the energy layers, spot positions, and weights are found through plan optimization.

Plan optimization

An IMPT treatment plan is composed of the list of machine parameters defining the treatment delivery: the beams (each characterized by a gantry and a couch angle) and the list of spots (x, y, E, w) for each beam. These parameters are optimized so the resulting dose agrees with the prescribed dose by the physician.

A list of objective functions is established to guide the optimization to converge to the most optimal dose distribution. The desired absorbed-dose levels for the delineated structures (e.g., target volumes and OAR) are specified in several optimization objectives. Typically, objectives are based on target volumes and OAR using minimum, maximum, and average dose metrics or dose-volume histogram (DVH) metrics. The DVH constraints can be formulated as “at least 98% of the prescribed dose to ‘x%’ of a volume” (e.g., D98% or Dnear-minimum dose) or “not more than 105% of the prescribed dose to an absolute volume” (e.g., D0.03 cm^3 or Dnear-maximum dose). For the OARs in parallel, mean doses or

the percentage of volume receiving a certain dose are specified (e.g., V20Gy). On the other hand, for serial OAR, maximum doses are used, such as 0.03 cm^3 of the spinal cord should receive less than 45 Gy (e.g., D0.03 cm^3 (spinal cord) < 45 Gy). Weight factors are given to each objective to reflect their relative importance. Therefore, the full objective function that will be optimized corresponds to a weighted sum combining all the objectives. Iteratively, the objectives and weights are manually modified to achieve the best plan (i.e., the highest dose to the tumor and the lowest dose to OAR) following the physician's feedback.

Robust optimization is the state-of-the-art for IMPT planning since it has shown superiority over margins. In robust optimization, the planning is done directly on the CTV, and the dose depends on a set of uncertainty parameters. Several robust optimization methods have been proposed in the literature [74, 97, 98, 99]. In this thesis, the worst-case robust optimization available in the commercial TPS RayStation was used to create all IMPT plans [72]. The worst-case scenario evaluates the full objective function in a group of uncertainty scenarios and minimizes the worst of them. Typically, optimization objectives are set as robust against setup and range errors. Including anatomical scenarios for moving targets is also possible (see Figure 1.10). From the uncertainties given as input, error scenarios are simulated and combined to generate a pool of uncertainty scenarios. During worst-case scenario optimization, dose distributions are computed in the background for the sampled error scenarios (e.g., 21 scenarios in Figure 1.10), penalizing the worst-case scenario and changing the plan parameters (e.g., spot weights) accordingly. In every iteration, the final computed dose gets closer to meeting the objectives set as robust not only in the nominal dose distribution (i.e., representing no uncertainty) but also in all error scenarios from the sampled group, including the worst-case.

1.4.5 Plan evaluation

In standard radiation therapy, the plan evaluation is based on the PTV coverage: if the PTV is properly covered, the plan is considered robust enough. However, this is not extensible for IMPT plans that were robustly optimized based on the CTV. The evaluation of a robustly optimized IMPT plan requires both the analysis of the nominal dose (i.e., final dose distribution without explicit simulation of errors) and the study of its behavior in the presence of uncertainties. The nominal dose shape is visually inspected via the isodose lines and evaluated according to the clinical goals. Later, the dose is evaluated in simulated uncertainty scenarios following what is known as robustness evaluation.

The main objective of the robust evaluation is to check if the robustly optimized plan against specific pre-defined uncertainty levels is indeed that robust. To do so,

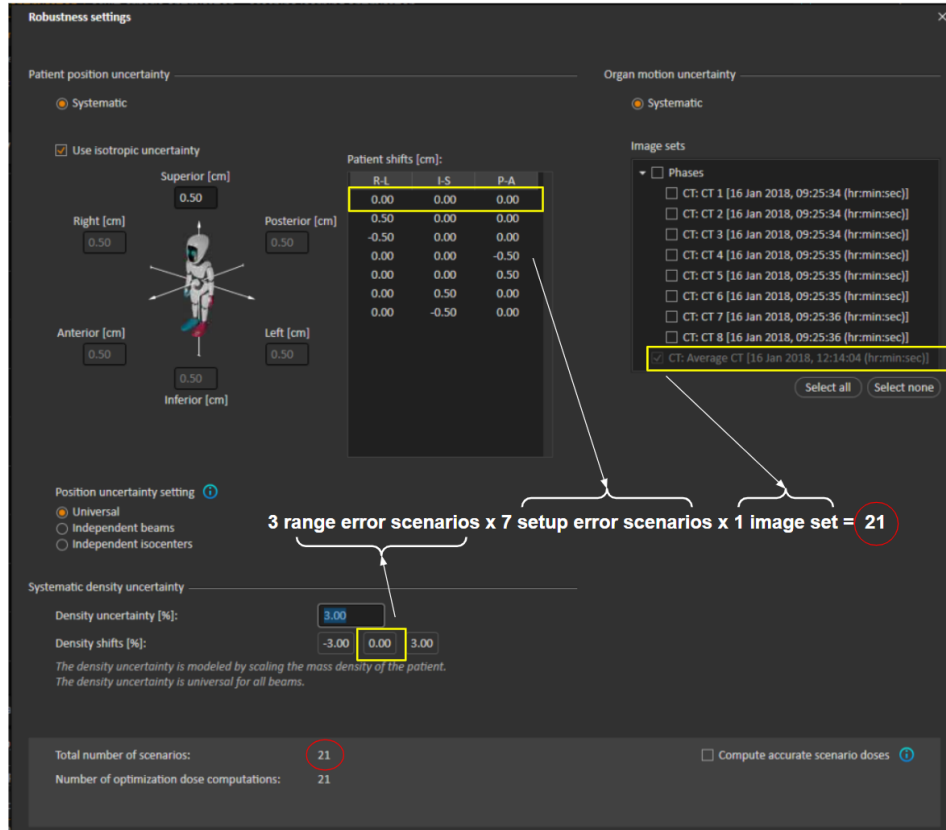


Figure 1.10: Robust optimization window from RayStation 11B. The nominal scenarios (i.e., with no error) for each uncertainty are highlighted in yellow. The combination of 7 setup uncertainty scenarios, 3 range error scenarios, and a planning image results in a robust optimization using 21 error scenarios

the nominal dose is recomputed in a group of uncertainty scenarios at least similar (or including addition error scenarios) to the ones used during optimization. The DVHs corresponding to the dose distributions from each of the simulated error scenarios are displayed all together as DVH-bands (see Figure 1.11). The thinner the DVH-band, the more robust the plan. A plan is considered robust enough when the set of clinical goals is achieved not only in the nominal scenario but also in the sampled error scenarios. However, the width of the DVH-bands and the confidence levels depend on the method chosen to sample the scenarios [100]. Korevaar et al. suggested an evaluation strategy based on voxel-wise minimum and maximum scenarios [101]. Voxel-wise minimum scenario is constructed by taking the minimum value independently for each voxel over all the sampled scenarios. Such an approach evaluates the dose in not physically realizable (i.e., unrealistic)

scenarios providing stringent evaluation thresholds. Moreover, the limited range of possible simulated scenarios in commercial TPS difficulties an accurate estimation of plan robustness. To overcome this issue, Sterpin et al. addressed such limitations and proposed a statistically sound strategy where scenarios were sampled according to their probability [100]. This alternative may provide an improved insight into the actual robustness of the plan for a given confidence level. Theoretically, the robustness of the plan is directly linked to the need for plan adaptation during the treatment course. The more robust a treatment is, the less need for plan adaptation. Therefore, an accurate estimate is crucial to find a good compromise between both parameters according to the clinical needs.

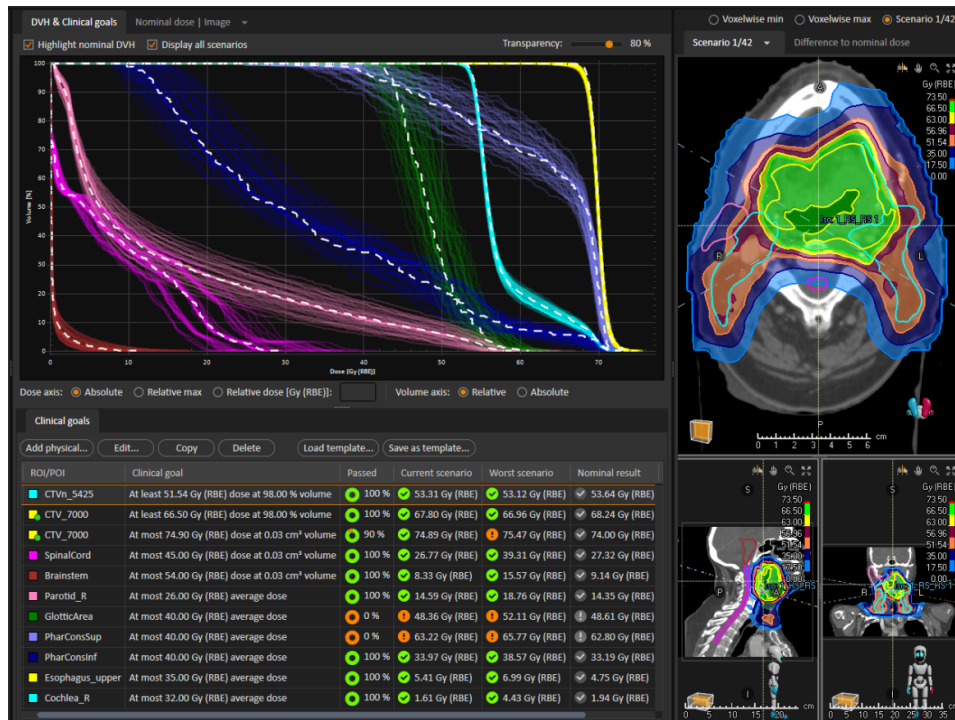


Figure 1.11: Robust evaluation window from RayStation TPS. For a better visualization only few structures were selected. On the top left the DVH-bands are display, each line corresponds to a simulated uncertainty scenarios. The DVHs from the nominal dose distribution (i.e with no uncertainty) are highlighted with the white dashed line. On the bottom left, the values for the clinical goals in the nominal, worst-case scenario and current scenario (first scenario from the 42 explored scenarios) are displayed. On the right appears the navigation window to explored the dose distributions from all simulated scenarios.

1.4.6 Quality Assurance

Proton accelerators are very complex devices built with many different parts. Therefore, errors can occur when delivering the treatment plan. Patient-specific quality assurance (QA) protocols are designed to verify that the IMPT plan generated on the TPS correctly delivers the intended dose distribution. Patient-specific QA is conventionally carried out with measurements on phantom-based systems. An increased interest has been shown in alternative QA tools: machine control and log files. These QA tools can help to speed up the QA process and have been shown to be more sensitive to data transfer, and delivery errors than patient-specific QA measurements [102, 103]. Machine control files contain the information sent to the delivery machine and validate the data transfer. Log files are historical records of each dose delivery, including information such as gantry positions, delivered dose rate, spot positions, and MU output as a function of time. Delivered dose distributions can be reconstructed using these log files based on the delivery's machine parameters. The reconstructed dose can then be compared to the clinically planned dose. Therefore they can verify both data transfer and dose delivery [102, 104].

1.4.7 Treatment delivery

Once the patient-specific QA is passed successfully, the patient is ready to receive the treatment in several fractions. For an accurate treatment delivery, the treatment couch and the patient must be in the same position as in the pCT image where the treatment plan was generated. With this purpose, immobilization devices have been historically used for patient alignment. Nowadays, in-room imaging is additionally performed to improve patient setup guidance [105]. Images are often acquired using embedded systems on the treatment machines, as portal imaging devices allowing for 2D orthogonal X-ray images or CBCT images providing 3D patient imaging. Dedicated in-room CTs are also available in some treatment rooms but not widely available yet [106]. The image of the day is typically aligned with the pCT using RIR algorithms. According to the registration information, the treatment couch is rotated and translated until the initial patient position is correctly reproduced. Moreover, daily imaging can provide information on the anatomical changes in the patient. Although imaging techniques are improving, some limitations can still cause residual setup errors. These setup errors can arise from constraints of the imaging system (e.g., noise, resolution or artifacts), misalignments between the imaging and delivery systems, uncertainties in registration algorithms, or anatomical changes.

When the patient is correctly re-positioned, the treatment fraction can be delivered. At this point, strategies to fight against intra-fractional uncertainties are put in place. For instance, breathing motion mitigation methods such as deep inspiration breath hold (DIBH) or mechanical assisted non-invasive ventilation (MANIV). DIBH has shown a protective effect and dosimetric benefits in heart and lung sparing [107, 108] for IMPT treatments. Although MANIV has not been applied to IMPT, it can reduce inter-fractional tumor motion amplitude and stabilize the respiratory rate, which may facilitate a more accurate PBS delivery [109, 110]. More sophisticated methods like gating or tracking have also been proposed [111, 112]. The latter requires complex 4D delivery to synchronize the plan to the actual patient motion during treatment, which is difficult to implement clinically. Additionally, more advanced imaging techniques have been developed to assess in-vivo range uncertainty before, during, and after delivery [113]. It is the case of proton radiography [114] or prompt gamma cameras [115].

As reviewed in this section, much effort is put into handling treatment uncertainties during treatment planning and delivery. Unfortunately, treatment plans have the limitation of relying on the pCT, assuming the patient's anatomy will remain constant during treatment. However, patients are not static, and variations can and do occur. What is planned, even when accounting for some of the uncertainties, is unfortunately not always what it is delivered. Therefore, APT remains necessary in some cases to compensate for the dose distortions induced by anatomical changes [22, 80].

1.5 Adaptive radiation therapy

As stated in the seminar paper by Sonke et al., "Adaptive radiotherapy aims to deliver radiation accurately and precisely in the presence of anatomical changes. To that end, ART uses an imaging feedback loop to quantify these changes and modify the treatment plan accordingly" [20]. The choice to clinically implement ART in a particular time scale and treatment location is based on the clinical goal, the available technology, and the human and material resources. In this section, we will address: (1) types of anatomical changes, (2) the rationale for online adaptation specifically for proton therapy, and (3) the steps of a standard online adaptive workflow following the current clinical practice.

1.5.1 Anatomical changes

IMPT dose distributions are likely to suffer dose distortions due to their sensitivity to changes in the patient's anatomy [30]. An excellent example to illustrate the problem of IMPT dose distortions is weight loss, which is very common in

patients under radiation therapy treatment [116, 117]. In the presence of weight loss, the initially planned proton beam will go through a reduced amount of body tissue and a more significant amount of air. Therefore, the planned positions of the Bragg peaks (which depend on the traversed tissue) will be modified, causing dose distortions [118, 21]. For a better understanding of concepts introduced later in this thesis, anatomical changes will be divided into density changes and non-rigid anatomical changes.

Density changes

When density changes occur, the patient's anatomy can experience residual rigid translations, but it remains mainly in the same position with an increased or decreased density in a specific region. For instance, nasal cavity filling illustrate how the air can be replaced by mucus, leading to an increased density inside the cavity as shown in Figure 1.12.a. This kind of density changes can also occur in the thoracic region. For instance, the air in the lungs can be replaced by a fluid due to pleural effusion or by tissue when atelectasis occurs. Another example of rigid density changes could be tumor shrinkage without involving modification of the surrounding tissues (Figure 1.12.b). These kind of anatomical changes are usually not a problem in photon therapy where the *stable dose cloud approximation* can be assumed but can induce critical dose distortions in proton therapy due to range errors.

Non-rigid anatomical changes

Non-rigid anatomical changes are characterized by a modified relative position between the OARs and the target volumes. For instance, tumor shrinkages, where the free volume liberated by the tumor is replaced by an OAR that enters the prescription dose region. More complex reorganizations of the anatomy, such as reconfiguration of the anatomy within the abdomen [119], baseline variation in the lung [120, 121, 122] or soft tissue modifications in the head and neck region [123] are also good examples of non-rigid anatomical changes. Since these kind of changes in the patient geometry can displace the target volume out of the dose prescription region they can benefit from plan adaptation in both photon and proton therapy.

1.5.2 The rationale for online adaptive proton therapy

The sensitivity of proton doses to anatomical modifications entails a higher need for treatment plan adaptation to ensure an appropriate dose delivery. Hoffman et al. found that for the same patient cohort of 23 lung cancer patients, 61% needed adaptation when treated with protons to ensure target coverage, while no

Dose distortions of IMPT plans caused by density changes

a) Cavity filling

b) Lung tumor shrinkage

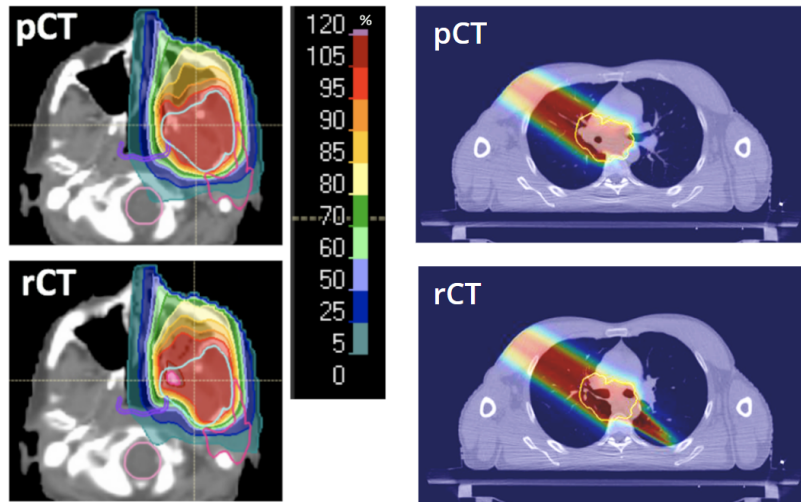


Figure 1.12: Effect of different anatomical variations causing dose distortion in various IMPT plans. The cavity filling example was extracted from Bernatowicz et al. [33]), and the lung images are courtesy of Guillaume Janssens from IBA.

patients required adaptation when treated with photons [22]. Robustly optimized IMPT plans have improved robustness compared to margin-based plans [77], providing both superior target coverage and OAR sparing in the presence of various treatment uncertainties and partially reducing the need for adaptation [79, 123]. The added value of including multiple-CT for anatomical robust optimization has also been investigated for HNC [124, 125, 126] and lung cancer [127] to increase even more the robustness of IMPT plans. Nevertheless, there are unforeseeable anatomical changes related to treatment delivery that are difficult to predict and simulate, such as tumor shrinkage or patient changes in weight or shape. Therefore, treatment plan adaptation remains necessary to compensate for anatomical changes that the robustness of the initial plan fails to account for.

Fraction after fraction, these anatomical changes can be monitored to trigger the need for adaptation when necessary [81, 128]. Hu et al. retrospectively analyzed robustly optimized IMPT plans for 711 patients in approximately weekly repeated CTs (rCTs) and reported adaptation rates for different locations [26]. From all the explored treatment sites, HNC and lung or chest diseases have the

highest re-planning rates with 54% and 58% , respectively. However, the rate for offline adaptation varies a lot across studies, planning strategies, locations, and probably clinical experience [25]. For instance, Jíří et al. [129] reported a very high adaptation rate, with almost all nasopharyngeal cancer patients requiring treatment plan adaptation. On the contrary, Scandurra et al. [123] found a much lower adaptation rate for the same location with only 2 out 25 patients requiring plan adaptation. In lung cancer patients, the need for plan adaptation has been proven necessary in clinical experience [130], or even mandatory [22]. Koay et al. reported that from a lung cancer patient cohort of 44 patients, 9 required adaptation, particularly those with large tumors that shrank substantially during therapy. They also concluded that APT could reduce doses to normal tissues and prevent target misses in the 2 cases that would have been compromised without adaptation [69].

With the currently available technology, APT is typically performed offline in clinical practice. In an offline setting, a few days can pass from acquiring a new CT image to delivering the clinically approved adapted plan. Offline ART also has the drawback of being particularly labor-intensive and, therefore, represents a significant personnel and material cost. Although improving the delivered dose, especially in case of slow inter-fractional changes, offline APT is inadequate in the presence of faster (daily) changes occurring from one fraction to the next. For instance, reconfiguration of the anatomy tumor response [131, 132, 133], weight loss [117], nasal cavity filling [33, 124], or quite complex tissue variations in shape and position of the cervix-uterus [134]. The slowness and the need for additional resources inherent to offline APT hamper a more recurrent (e.g., daily) adaptation that could adequately address these inter-fractional changes. Alternatively, a fully automated online adaptive approach needing fewer resources could be very well-suited to immediately address such inter-fractional changes that sometimes can be fast and unpredictable.

As a result of all of the above, there has been growing interest in online adaptation workflows, especially for proton therapy, since the increased need for treatment plan adaptation adds a significant workload when using offline APT [135, 81]. However, the clinical implementation of online APT is challenging. From the moment the adaptation is judged as required to the performance of quality assurance of the re-optimized plan, the patient remains on the treatment couch at the treatment position. Therefore, a successful online adaptation workflow needs to be extremely fast, within some minutes. This process requires a high level of automation and integration to ensure an efficient workflow. In recent years well-integrated systems, such as MR-Linacs and the Ethos System (see section 1.7.1), have been developed specifically for online ART, improving the efficiency of this

process [136, 137]. However, online APT has not been clinically implemented yet.

Online APT could optimally manage inter-fractional changes while tackling the issues introduced in this section and summarized in the following lines. First, the use of margins and robust optimization does not account for all anatomic changes over the treatment course. Secondly, offline adaptation is the most common clinical practice to compensate for anatomical changes but remains a very tedious process. Third, although increasing the robustness of the initial treatment plan could reduce the need for offline adaptation, robust optimization comes at the cost of increasing the dose to the surrounding organs to account for uncertainties. Such an approach would therefore undermine the dose conformity associated with IMPT treatments. Forth, to fully profit from the adaptive approach, the whole process needs to be established within the time span of the anatomy changes, which sometimes occur from day-to-day being the offline approach too slow to follow. Therefore, there is still a clinical need for daily dose analysis and access to adapted plans to compensate for dose distortions that could compromise target coverage or OAR sparing. The application of online APT on a daily basis (i.e., at all treatment fractions) is known as daily APT (DAPT). Finally, besides ensuring that the patient receives the intended dose, online (D)APT has multiple benefits that could additionally improve the patient outcome of IMPT treatments:

- **Reduction of treatment uncertainties:** Due to the re-optimization in a repeated image, the errors in the delivered dose due to anatomical changes can be corrected, and the positioning uncertainties can be reduced.
- **Potential for more conformal dose distributions:** Because the uncertainties are reduced, more conformal planning approaches, which would be anatomically un-robust if delivered without adaption, can be explored (e.g., less anatomical-robust planning or smaller CTV-PTV margins). These new treatment planning options could enable a higher sparing of the OARs and reduce toxicity.
- **More accurate estimation of the delivered dose:** Since dose calculation is performed in an image representing the actual patient geometry just prior to treatment delivery, the adapted dose results in a more accurate estimate of the actual dose delivered to the patient. This is relevant for both reporting and monitoring the treatment.

1.5.3 Adaptive radiation therapy workflow

The flowchart in Figure 1.13 illustrates a standard adaptive workflow, including online and offline strategies. Although both workflows contain almost the same steps, each adaptation branch uses a different planning image and occurs in different time scales. Online adaptation should be completed in a few minutes, while several hours or even a few days fit the offline adaptation time frame. Therefore, online workflows have particularities that speed up the process compared to the standard radiation therapy planning described in section 1.4.

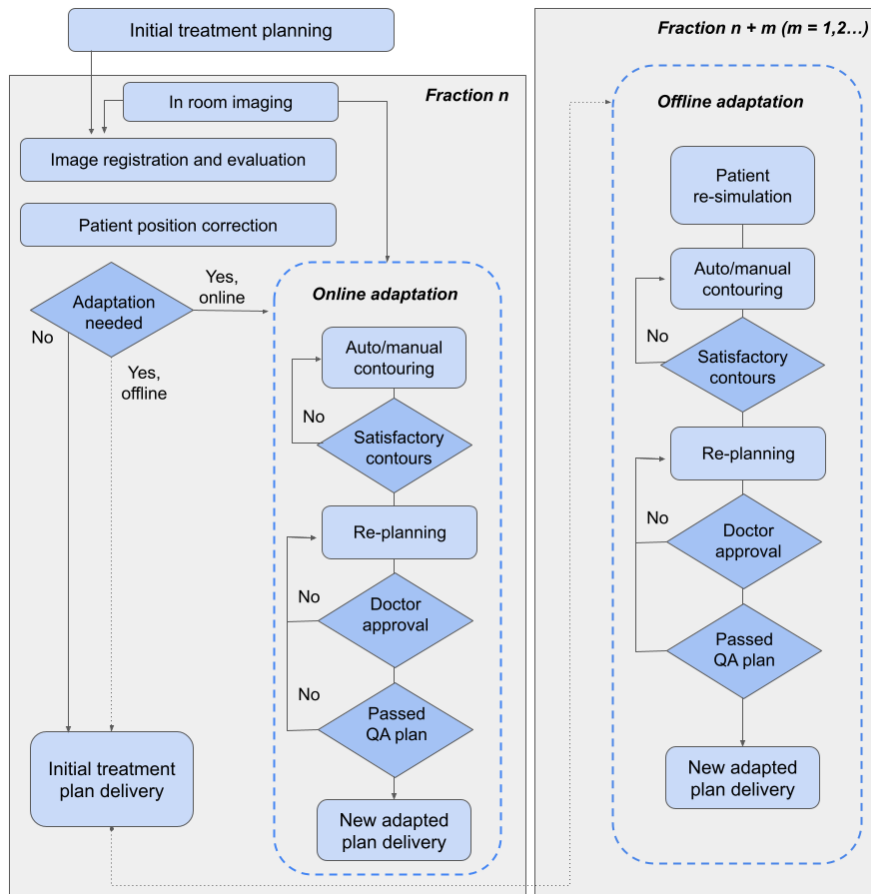


Figure 1.13: Flowchart of standard treatment plan adaptation workflows for online and offline strategies. Diamonds represent steps requiring decision-making. Adapted from [138]

Patient re-positioning and re-imaging

Accurate patient imaging to quantify the patient's anatomy at each fraction is the first step of any online ART workflow. Imaging can be performed with an embedded system in the treatment machine, such as a CBCT, or with an in-room independent system like a CT on-rails. Rigid registration between the image of the day and pCT is often used for patient re-positioning to match the isocenter between the two images. MRI proton therapy will be disregarded in this section since, although under research, it is not clinically available yet.

The higher image quality is an essential advantage of CT on-rails over CBCT imaging. Accurate HU information is crucial for online APT since slight HU differences may introduce range errors and inaccurate final dose calculation. CT images are also more suitable than CBCT images for contouring. In terms of the extra whole-body dose over the course of the treatment, CT-on-rails might give a slightly lower dose than CBCT [106]. Moreover, low-dose scanning protocols could provide even lower doses with acceptable image quality [139]. Nesteruk et al. proposed CT-on-rails as a suitable modality for online APT since the expected uncertainties did not affect the dosimetric efficacy of their adapted plans [106].

Re-planning based on CBCT images also has great interest due to the low cost, compactness, and thus wide availability of this imaging system. CBCT images are known for their lower image quality and artifacts (e.g., noise, beam hardening, aliasing, motion, or ring artifacts). In the past years, the increased interest in the clinical implementation of online ART pushed for the improvement in CBCT image quantification of commercial solutions. Moreover, analytical methods applied to CBCTs (e.g., scattering correction algorithms [140] or histogram matching [141]), DIR algorithms [142] and more recently, AI tools (e.g., neural networks further discussed in section 1.6.2) have been explored and compared for synthetic CT (sCT) image generation and proton dose calculation [143, 144].

Contour definition

As described in section 1.4, contoured volumes are needed to define the objective functions given to the optimizer. In offline APT, volume contouring is performed manually on rCTs. However, in online APT, contour generation should be as fast as possible. Therefore, automatic tools can be used to map the initial contours to repeated images using rigid or deformable registrations. AI tools will also play a vital role in this step, as the Ethos System has already shown. Today, contour generation is still considered the bottleneck of ART and APT workflows due to the required manual verification (and correction) of automatically generated contours.

Assessment

A key component of offline and online APT workflows is assessing the adaptation need. It is often unclear what the appropriate frequency of adaptation should be and whether plan adaptation is even needed from a dosimetric, TCP, or NTCP point of view [145]. Figure 1.14 shows three adaptive workflows. Nowadays, offline adaptation is often made at fixed intervals where rCTs are acquired (e.g., halfway through the fractionated treatment or weekly) [25]. The clinical implementation of online APT could enable plan adaptation triggered at any point of the treatment course by observed changes resulting in a negative effect on the plan quality. Frequent online adaptation (e.g., daily) could also be implemented if resources permit.

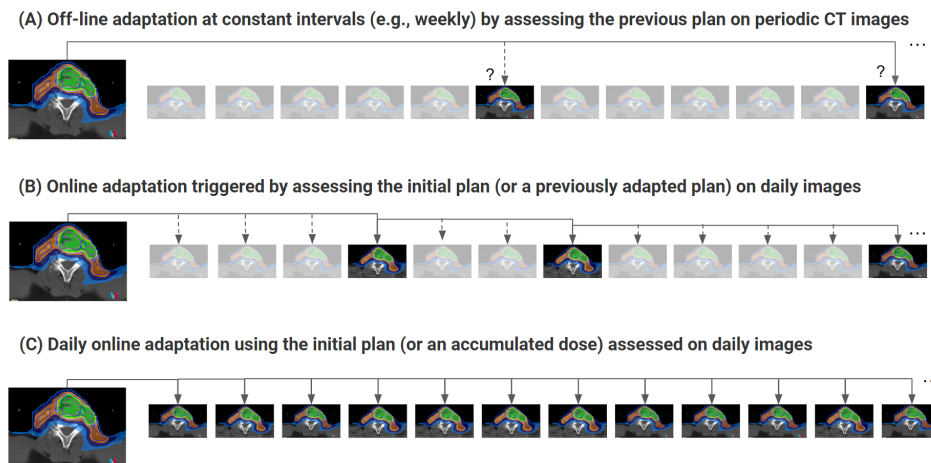


Figure 1.14: Different adaptation workflows. Figure based on Heukelom and Fuller [146]

The most simple decision-making protocol is to re-compute the initial treatment plan on the repeated image (i.e., the dose intended to be delivered) and then use the new generated contours for evaluation. If treatment plan adaptation has already occurred during the treatment course, the last adapted plan could be used instead. More sophisticated approaches suggest using DIR to estimate the cumulative dose over all previous treatment fractions and judge the need for adaptation based on the summed dose [147, 148]. This accumulated dose estimates the total dose delivered to the patient until that moment (Figure 1.15). For an even more informed decision, log-file dose reconstruction could also be included in the loop to better evaluate the delivered dose to the patient [149].

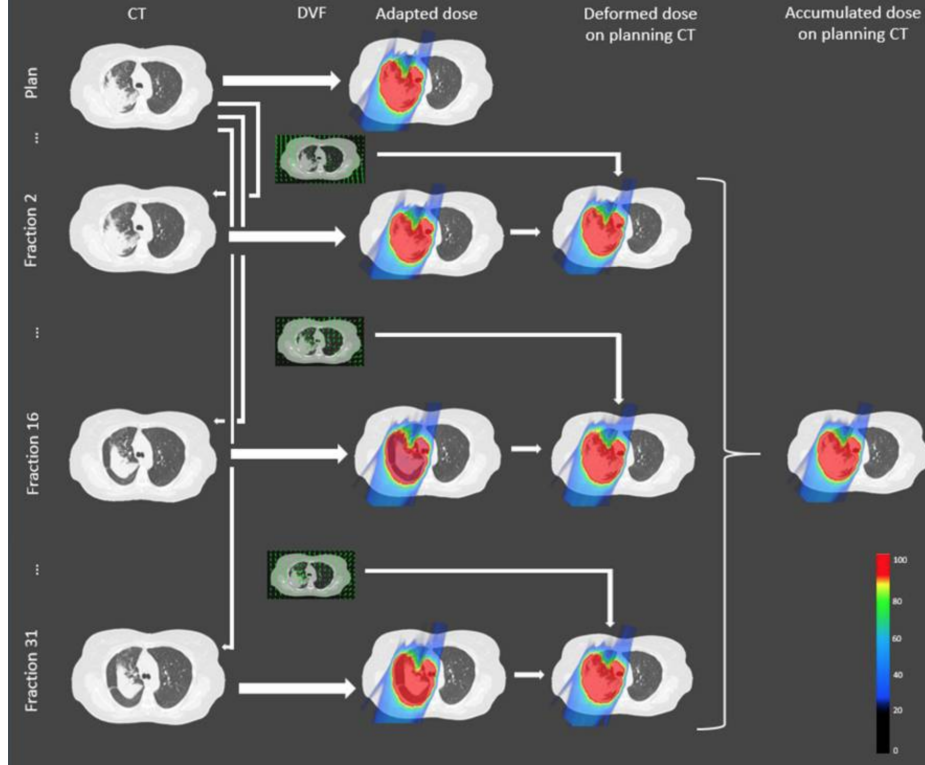


Figure 1.15: An example of the dose accumulation workflow using the planning-CT (pCT) as the reference image. Deformable vector fields (DVF) for three fractions are displayed. Adapted dose distributions on the repeated-CTs and their deformation to the pCT are also shown for these fractions. Tumor shrinkage can be observed on the repeated-CTs. From [135]

Re-planning and re-optimization

The re-planning step aims at generating a new dose distribution according to the new patient anatomy. For IMPT plans, this means a re-optimization to calculate a new plan containing updated energy layers, spot positions, and spot weights (x, y, E, w) . In an integrated APT solution, this process would use the optimization and dose calculation algorithms already embedded in commercial TPS. Online re-planning strategies are the main focus of this thesis and they will be further explained in the following chapters. In Chapters 2 and 3 a dose restoration method is evaluated in lung and head and neck locations, respectively. More complex re-planning strategies, including ML dose prediction (see section 1.6.4) are explored in Chapters 4 and 5.

Plan evaluation and doctor approval

Once the final adapted dose distribution is available, the physician's approval is needed to continue the workflow until treatment delivery. Since the presence of the doctor is required in the treatment room, plan approval is also one of the bottlenecks of the online adaptive workflows, together with manual re-contouring (or adjustment) of the new anatomy.

Quality assurance and treatment delivery

Quality assurance of the approved plan is the step that follows. A key challenge for online ART, is the inability to perform the conventional patient-specific QA measurements (e.g., to deliver the plan to a phantom) as the patient remains on the treatment couch. Secondary and independent MC dose calculations on the new patient anatomy have been proposed instead to address the new plan verification [150, 102, 151, 152, 153]. The use of machine delivery log-files after dose delivery has also been suggested. Moreover, Matter et al. showed log-files to have higher sensitivity compared to patient-specific verification measurements [103]. However, when using log-files, QA is performed after dose delivery [149, 154], meaning that correction can only occur in the next fraction.

1.6 The role of artificial intelligence in online adaptation

Challenges for clinical implementation of any adaptation strategy are: increased workload [146], longer treatment time per fraction [155], limited image quality [135], radiotherapy technologists (RTT) training [156, 157], dose accumulation uncertainty [158], and software or workflow implementation [159, 160]. Solutions to many of these challenges can be found through task automation, a very well-suited application for AI. Moreover, manual processes in the radiation therapy workflow are affected by high inter-planner variability, which may result in sub-optimal solutions. In this section, we will introduce the main concepts related to AI that will be mentioned at some point in the manuscript.

1.6.1 A brief introduction to neural networks

AI is the term used to refer to the methods and algorithms aiming at reproducing human actions requiring specific knowledge learned through human experience. As explained in Barragán-Montero et al. [161], Machine learning (ML) is part of the branch known as "Connectionism" of the whole AI concept. ML has two main phases, which occur one after the other, training and inference. During training,

previously acquired data is used to extract meaningful features and patterns. During inference, these parameters are applied to new data (not seen during training) to perform a prediction, classification, or another decision-making process. When sophisticated hierarchical structures were introduced in the training process, the term Deep Learning (DL) was coined.

ML algorithms can only deal with well-structured and labeled data. On the contrary, DL algorithms can learn from data without manual human intervention to define relevant and specific features (i.e. unlabeled data). DL imitates the way the human brain works by creating an artificial network of neurons. An artificial neural network (ANN) consists of many interconnected layers of neurons. The mathematical model of a single neuron is very similar to the biological one. A group of inputs (dendrites) is multiplied by specific weights (synapsis). The weighted inputs are then summed together following the shape of a non-linear activation function generating an output (axon) that might be used as the input (dendrite) of the next mathematical structure (neuron). The first and last layers are the input and output, respectively. In between these two, one or several hidden layers can be interconnected. The more neurons the hidden layer contains, the more complex functions can be approximated. The processes that occur at high complexity levels in the hidden layers are why the technology cannot be easily interpreted and is therefore referred to as a black box.

Convolutional neural networks (CNNs) are a type of ANN used in medical imaging since they excel at processing data with a grid-like topology, such as an image composed of pixels. They are known as 'convolutional' networks as they employ a mathematical operation known as convolution. UNets, are a special type of CNN characterized by a U-shape consisting of a contracting path (first part of the U) and an expansive path (second part of the U). They were first introduced in 2015 by Ronnenberg et al. for Biomedical Image Segmentation [162] and have become incredibly popular since then.

Complex architectures, such as GANs (Generative Adversarial Networks), combine more than one network. GANs are composed of a generator and a discriminator: the generator learns to generate fake but realistic (plausible) data and the discriminator learns to distinguish the generator's fake data from real data. When training starts, evident fake data is produced by the generator. The first generated data can be classified without a hitch as fake by the discriminator. However, as the training progresses, the generator is penalized for producing implausible results and gets better at generating realistic outputs that are not easily recognized by the discriminator. If the training is successful, the generator can produce such realistic data that the discriminator can not distinguish between real and fake images.

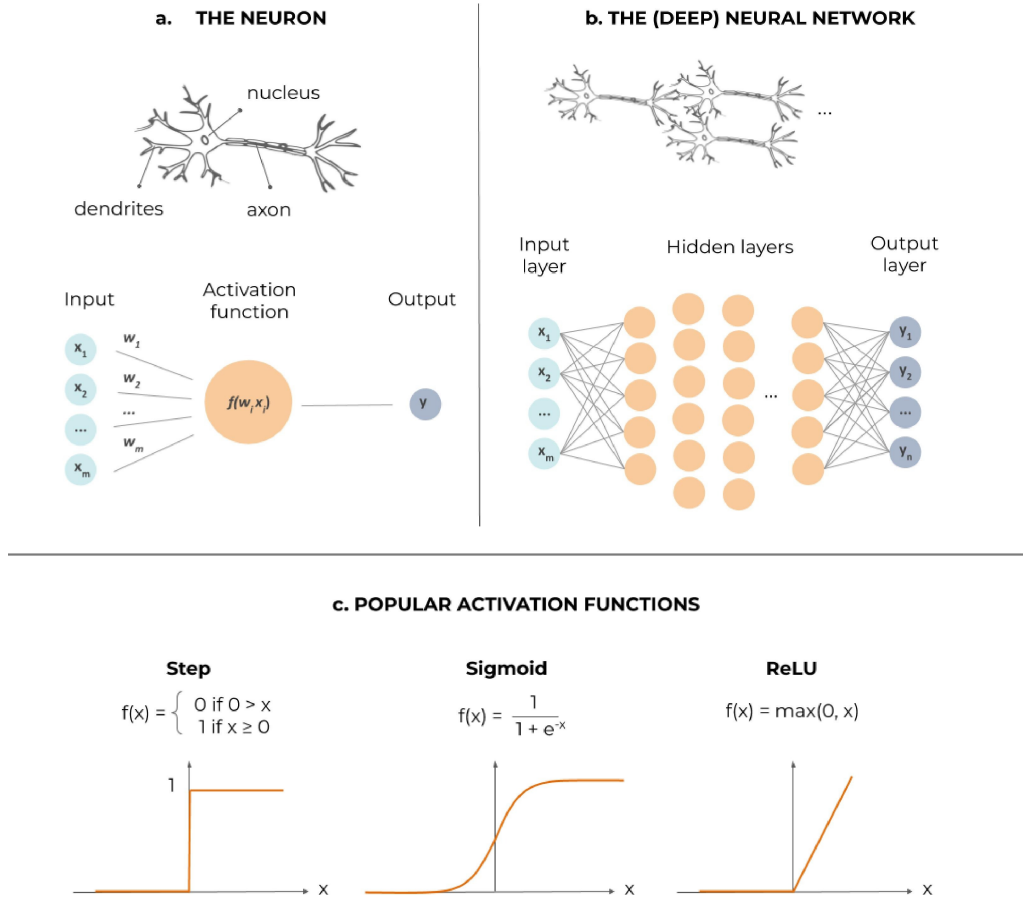


Figure 1.16: Artificial neural explained with mathematical-biological parallelism. a) A neuron, processing several dendritic inputs through a nonlinear activation function f , to produce an axonal output. (b) The neurons can be interconnected in a feed-forward way, into successive layers; as soon as a nonlinear 'hidden' layer is inserted in between the inputs and outputs, the network can potentially approximate any function; specific activation functions can be fitted in the output layer to achieve either regression or classification tasks. (c) Examples of nonlinear activation functions in the hidden layers: the step function, from biological inspiration, the sigmoid, its continuous and differentiable surrogate, and the rectified linear unit (ReLU), that improves training of deep layers. Courtesy of A. Barragán-Montero from the review "Artificial intelligence and machine learning for medical imaging: A technology review" [161]

Although AI tools have explored a vast number of applications in the radiation therapy field, we will focus on three key steps of the online adaptive workflow that can enormously benefit from the consistency, rapidity, and automation offered by AI-based solutions: sCT generation, automatic segmentation, and automatic treatment planning.

1.6.2 Synthetic CT generation

Since the attenuation coefficients encoded in the CT image are needed for dose calculation, the development of techniques that can generate sCT information from other modalities is required to implement online adaptation based on daily CBCTs or MR-only radiotherapy.

Most of the commercial solutions clinically used for online ART are based on sCT generated by DIR. However, low performance of the deformation algorithms can occur when new elements that were not present on the pCT (that will be deformed) appear on the repeated image, such as air pockets. The direct sCT generation from an image containing the real patient anatomy should reproduce more accurately the patient's geometry of the day. Several AI tools have been recently explored, with GANs being extremely popular since unpaired data from the different modalities (MRI, CBCT and CT) can be used for training [163, 164, 165, 166, 167, 168]. Although paired data is not easy to obtain, other architectures like UNets have successfully generated sCT using paired CT-CBCT data acquired on the same day or very close in time [169, 170]. A recent study from Thummerer et al. demonstrated that sCTs generated from MRIs and CBCTs using a deep CNN were both suitable for daily IMPT dose calculation [171].

1.6.3 Automatic segmentation

Volume segmentation can be easily modeled as a classification task in which each pixel is associated with a corresponding class (e.g., OAR or target volume). Therefore, many AI solutions have been proposed for volume segmentation in chest, brain or head and neck [172, 173, 174]. Head and neck is particularly challenging due to the small organs to segment (parotids, submandibular glands etc). In the same way, target volume was first applied to more straightforward problems like GTV delineation [175, 176, 177] or easy locations like lung and brain. However, it is much more challenging in other regions since CTVs are often generated using guidelines and explicitly adjusted for each patient with the surrounding anatomical structures. More and more studies are addressing target delineation, and nodal regions in HNC [178, 179, 180].

Automatic contouring is today the most used ML application in the clinic. It is already available in commercial solutions [181], and also in-house generated models have been implemented in different clinics to ease the cumbersome task of volume contouring [182]. However, ML methods for automatic segmentation still lack QA tools to assess their interpretability and risk of failure, and they can not be trusted blindly. Today, the visual inspection of a medical expert, who edits the contour in the regions when necessary, is still required. Therefore, the process

systematically requires the presence of a medical doctor in the room. Nevertheless, AI tools can still provide precious time savings as the 33% time reduction compared to fully manual HNC segmentation reported by Van der Veen et al. [183].

1.6.4 Automatic planning

Much effort has been put into developing and exploring strategies for automatic planning in radiation therapy, from planning libraries (e.g., RapidPlan from Varian Medical Systems) to more sophisticated methods based on deep learning tools to predict 3D dose distributions [184]. In this thesis, we will explore a two-stage knowledge-based planning workflow using a predicted 3D dose distribution followed by an optimization algorithm known as dose mimicking [35]. This automatic planning workflow is described in Figure 1.17.

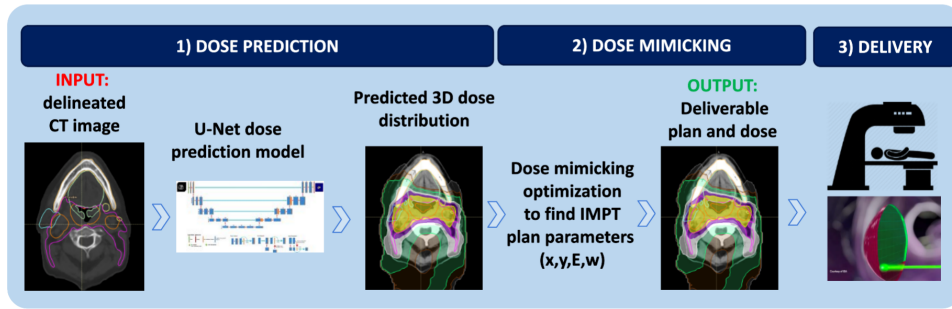


Figure 1.17: An overview of a complete knowledge-based planning pipeline combining (1) dose prediction and (2) dose mimicking optimization.

In the first step, an ML model receives as input the pCT and the delineated contours to predict the desired dose for each voxel of the patient geometry. The model's output is a 3D dose distribution that is not deliverable without the plan information. The goal of the second step (i.e., the dose mimicking optimization) is to find the set of machine parameters (x, y, E, w) in the case of IMPT) using the information stored in the predicted dose. At the end of the process, the output is an automatically generated deliverable dose distribution, which should be reasonably close to the predicted dose in the first stage. Many different neural network architectures can be used for dose prediction and there are also different ways of running the dose-mimicking step. Therefore, researchers have been trying to evaluate the output of several combinations of prediction and optimization models in photons [36]. However, fully automated workflows combining dose prediction and dose mimicking for IMPT planning have received less attention. The high complexity and variability of the information inside an IMPT plan and

the necessity of a robust plan generation make automatic planning for protons even more challenging.

1.7 State of the art in online adaptation

1.7.1 Online Adaptive Radiation Therapy (ART)

Until recently, the technical and logistical challenges of online ART made it practically infeasible for most radiotherapy centers. The introduction of AI [185, 186, 187] and graphics processing unit (GPU) based calculation engines [188, 189, 190] have speed up some of the steps in the adaptive workflow allowing to be performed them in a timeframe compatible with an online setting. As shown in the previous section, research is still ongoing to ease all the steps in the adaptive workflow. This section will explore the two most popular commercial solutions available for online adaptation for photons: magnetic-resonance linear accelerators (MR-Linacs) and the Ethos machine from Varian Medical Systems (Palo Alto, CA). For both systems, the presence of a doctor is still required in the treatment room.

MR-Linacs

The radiation delivery on the MR-Linacs is fully integrated with the MRI. This means the system can deliver treatment radiation beams and monitor the target area simultaneously using a non-ionizing imaging system. Therefore, physicians have better control over radiation delivery because they can see the internal anatomy and tumor with superior high-definition image quality, especially for some soft tissue cancers.

MR-Linacs allow for plan adaptation optimized on the daily MRI and contours. The first step is the registration between the pCT and the online planning MRI. The pre-treatment contours are automatically propagated by deformable registration onto the online planning MRI. If necessary, contours are edited by a radiation oncologist. The recontouring and correction of the contours is one of the bottlenecks of the workflow and can be translated in fractions that take up to 50 minutes to one hour to be delivered [191]. The plan is then recalculated or reoptimized on the online planning MRI and adjusted contours. Optimization is performed based on the pre-treatment planning objectives.

The daily treatment plan adaptation at almost every fraction provides better dosimetric outcomes than no adaptation. Nienier et al. analyzed 50 patients adapted online using an MR-Linac system for five different tumor entities (liver,

lung, multiple abdominal lymph nodes, pancreas, and prostate). They concluded that all subgroups clearly benefited from online plan adaption in terms of improved PTV coverage. Moreover, for the liver, lung, and abdominal lymph node cases, a systematic improvement in GTV coverage was found. Largely decreased OAR doses in the pancreatic cancer subgroup could also be achieved [192]. Moreover, Christiansen et al. could demonstrate that based on online plan adaptation with an MR-Linac, they could expect a PTV reduction, a dose reduction on the OAR, and thereby a reduction in toxicity based on NTCP models. They observed a relative reduction in complication probability of 91% for acute gastro-intestinal toxicity, 42% for urinary voiding pain, 25% for urinary incontinence, and 12% for hematuria [193].

Ethos system

The Varian Ethos system was recently developed as a completely self-contained online adaptive solution, and has been reported to be capable of performing adaptive treatments within 15–20 min [194, 195, 137]. Using AI, Ethos automatically segments normal organ structures that strongly affect the target volume, the so-called influencers. These contours are adjusted by the user if necessary and used to create a structure-guided DIR between the pCT and the acquired CBCT. An elastic DIR is also created between the pCT and CBCT. The elastic DIR is used to both propagate non-influencer OARs and generate a sCT by deforming the pCT into the CBCT geometry. For target propagation either structure-guided DIR or elastic DIR can be chosen. Once the structures are approved, the treatment plan is generated based upon a predefined “planning directive” optimized to the anatomy of the day and referred to as the adaptive plan. The original treatment plan is recalculated on the sCT and referred to as the scheduled plan. The user selects either the scheduled or adaptive plan for the treatment. The plan selected receives pre-treatment calculation-based QA, and post-treatment delivery log file-based QA. A limitation of the CBCT-guided online ART is the lack of continuous intra-fraction monitoring. As an alternative, a second CBCT is usually acquired just prior to treatment delivery for position verification.

First studies have already reported the benefits of online adaptation with the Ethos System achieving higher PTV coverage in the pelvic regions [194, 137, 196]. Sibolt et al. reported 76% of influencers requiring no or minor edits, and the adapted plan preferred over the scheduled plan in 88% of fractions [194]. In line with this results but showing a slightly improvement, Byrne et al. reported 91% of influencers with no or minor edits (and 72% for the target) and a selection of the adapted plan over the schedule plan in 95% of the fractions [137]. Yoon et al. demonstrated the feasibility of online adaptive HNC in the Ethos system also

below 20 minutes, similarly to the times reported for the pelvic region. They also showed that the daily adapted plans effectively spare the OAR while maintaining a therapeutic dose to the CTV [195]. The correction of the influencers segmented by AI has been reported in all studies as the most time consuming step in the full online ART workflow.

1.7.2 Online Adaptive Proton Therapy (APT)

Although online adaptation in conventional radiation therapy is starting to be more and more used with the arrival of the MR-Linac or Ethos Systems, its implementation in the proton therapy routine is still not in clinical use. Although much can be learned from online adaptation in photon therapy, online proton therapy requires different approaches due to its dosimetric characteristics.

The explored online adaptive workflows for IMPT can be divided into three groups: re-planning strategies, re-optimization and dose restoration approaches, and library plan based methods. The first group consist in a re-planning on the repeated image using the same objectives (or with slight modifications). The objectives are inherited from the initial plan and the traditional planning pipeline executed as explored by Matter et al [197] and, focusing on paranasal cancer patients, by Nenoff et al. [198]. In the second group, changes to the original plans are performed, using re-optimization [199, 200] or dose restoration approaches [34, 201, 33]. All these studies include beamlet energy adaptation and weight re-optimization. Regarding the locations that have been explored only the study of Bernatowicz et al. contained one lung patient, while the rest have been mainly focused on the pelvic region [34, 201] or on HNC [200, 199, 198].

In order to fit into the online time limitations, the dose restoration strategies used analytical dose calculation algorithms with rigidly mapped contours [33, 34]. These strategies aimed at restoring the original dose distribution in rCTs. Jagt et al. (2017) developed an in-house method involving two steps: restoration of Bragg peak positions by energy corrections and re-optimization of beamlet weights to recover the reference plan. With the presented method applied to 10 prostate cancer patients, 96.3% of the fractions resulted in $V95\% \geq 98\%$ in the prostate CTV, while $V107\% \leq 2\%$ was fulfilled in up to 92.5% of the fractions in the same contour [34]. Bernatowicz et al. presented a workflow that considered rigidly translated contours from the pCT to rCTs on the 3 patients examples: a nasopharyngeal case with varying nasal fillings, an oropharyngeal and a 4D-robustly optimized lung case. They investigated 3 ways of minimizing the difference between the original and the restored plan, focusing on: DVH points, voxel-wise dose comparisons and isodose contours. All adaptation methods recovered $V95\% \geq 95\%$ and $V107\% \leq 2\%$ for the CTV on all repeat CTs, with

the isodose approach yielding the best dose distribution as compared to the one originally planned. Since the isodose restoration approach seems to be promising, one of the objectives will be to evaluate this method in larger patient cohorts of lung and HNC patients.

Jagt et al (2018) expanded their previous methodology assuming manually delineated contours were available and removing the limitation to rigid registration [201]. They also restricted the number of proton energy layers, implemented a reference point method (RPM) optimization and allowed up to two iterations of 2500 beamlets in addition to the original plan. In this method, once the energies of the beamlets are corrected and the random beamlets are added, the plan is optimized, which may remove unused beamlets and energy layers. Beamlet addition and optimization may be repeated for a second iteration. In this case, all of the 88 scans produced $V_{95\%} \geq 98\%$ in the true prostate CTV, while $V_{107\%} \leq 2\%$ was fulfilled in 94.3% of the fractions in the same contour.

The third group addresses plan adaptation using plan libraries. For 6 patients with cervical cancer proton therapy, Jagt et al. (2019) demonstrated a full recovery of target coverage in 23 rCTs using a 2-plan library to warm start an online re-optimization of the daily treatment plan [202]. A recent study by Oud et al. [203] in HNC proposed the use of a plan library containing robustly optimized plans with different parameters. The approach resulted in overall reduced NTCPs (xerostomia and dysphagia) with limited cost in target coverage compared to no adaptation. However, a library of plan approach is based on a limited number of plans (typically 2-5) covering specific motion ranges. Moreover, the scans underlying the library of plans are typically acquired on the same day or first week of treatment and thus do not capture the full complexity of motion present during treatment.

Only a few explored MC dose calculation combined with deformably mapped contours [199, 200]. For the strategies using either manual or deformed contours, the adapted plan is not only a restoration of a previous treatment plan but a generation of a completely new plan accounting for anatomical deformations. Regarding the robustness of the adapted plans only Bernatowicz et al. [33] used robust optimization to generate the adapted plans, while the rest relied on PTV margins or previously optimized plans for Oud et al. [203].

1.8 Objectives and contributions

This section explains in detail the contributions made during this thesis to address the gaps in this research field previously mentioned: the clinical implementation of online APT to fight against inter-fractional changes which can compromise an accurate treatment delivery. Therefore, this thesis aims at exploring the feasibility and performance of **robust, accurate, fast and automated adaptive strategies** that could be applied in an online setting to improve the outcome of proton therapy treatments. To this end, head and neck and lung tumors were selected as they both show challenging inter-fraction variations. Proving that adaptation is feasible and beneficial for these challenging tumor sites may imply suitability also for other sites. To achieve this goal the following research questions are addressed.

Can the dosimetric effects of density changes be compensated by restoring the initial treatment plan?

An IMPT treatment plan can be described by the set of parameters of the individual pencil-beams: location, energy and intensity (x, y, E, w) . Density changes along the beam paths affect the depth of the Bragg peaks, resulting in a distorted dose distribution with hot and cold spots. This suggests that the dose distribution could be restored to its initial state by running a fast re-optimization to adjust pencil-beam energies and positions. Therefore, we implemented and evaluated a simple dose restoration strategy based on this idea for lung (Chapter 2) and HNC (Chapter 3). The method starts by rigidly mapping the delineated contours from the pCT to the rCT. After this, a fast re-optimization of the spot positions and intensities is guided by the isodose information of the initial dose aiming at minimizing residual dose differences between the initial plan and the restored (i.e., adapted) plan. A total of 27 rCTs of 14 lung cancer patients and 60 rCTs of 10 HNC patients were available for evaluation. In summary, the first two chapters of this thesis focus on the implementation and evaluation of a dose restoration workflow aiming at generating fully automated robust IMPT plans on rCTs.

Can robustness against breathing motion be handled by dose restoration?

The treatment planning and adaptation of proton therapy are even more challenging in lung cancer since the respiratory motion introduces additional uncertainties. The feasibility of a fully automatized 4D-robust dose restoration is explored in Chapter 2 for a lung cancer database. The breathing motion uncertainty is included in the robust optimization used in the dose restoration workflow. A robustness evaluation on repeated 4D-CTs follows to accurately quantify the robustness of the restored plans.

How much can we gain? What is the benefit of online-adaptive IMPT compared to robust treatment planning and offline adaptation?

It is known that the higher adaptation rate in proton therapy compared to photons is linked to the special sensitivity of proton therapy to uncertainties. Therefore, the necessity of adaptation could be reduced by replicating the *stable dose cloud approximation*, which is key and only valid in photon therapy. This approximation states that the initial dose distribution remains the same even in the presence of treatment uncertainties. By extending this formalism to proton therapy using dose restoration of the initial treatment plan, the higher offline adaptation rate of this modality could be reduced to the same residual numbers associated with photon therapy. The more the offline adaptation rate is reduced, the better. Less offline adaptation would imply no delays in the treatment delivery for the patient and more resources for the radiation therapy department.

We first investigate the gain of these methods compared to a non-adaptive robust treatment planning approach. To this end, we generate robust treatment plans for lung and HNC patients and study the adaptation rate required when no active adaptation is applied (i.e., when relying on the robustness of the initial plan). From these experiments, the impact of the anatomical changes captured in the repeated images can be studied, and the resulting dose distortions quantified. Then, the need for adaptation when using a robustly optimized initial treatment plan can be determined. Later, the adaptation is performed using dose restoration. By evaluating the quality of the restored plans and comparing them to offline adapted plans, we can assess to what extent the stabilization of the dose cloud provided by dose restoration can compensate for the dose distortions and reduce the offline adaptation rate.

What are the limitations of the restored plans?

The dose restoration described in Chapters 2-3 restores an initial treatment plan by accounting for density variations along the pencil-beam paths while ignoring changes in the shape and relative positions of the anatomical structures (i.e., tumor and OAR). This is an intrinsic limitation of the dose restoration method is addressed in Chapter 3 for a dataset of HNC patients in which such non-rigid deformations are present.

Is robust IMPT treatment planning feasible combining deep learning dose prediction followed by dose mimicking?

An alternative to online dose restoration is a fully automated planning performed without any previous information about another initial treatment plan. This could

be achieved by using deep learning dose prediction, which would take a CT image and the CT contours as input. After having the predicted dose value for each voxel, a dose mimicking optimization would be run to find the plan parameters to deliver a mimicked dose similar to the predicted one. The feasibility of fully automated robust IMPT plans combining deep learning dose prediction and dose mimicking is explored in Chapter 4. A particular focus is put on the dose mimicking optimization step, being the key to ensuring the robustness and accuracy of the automatically generated plans. This is especially challenging in HNC due to the presence of many radio-sensitive OAR close to the target.

Can dose mimicking enable fully automated robust IMPT planning for online adaptive IMPT?

The automatic IMPT methodology explored in Chapter 4 could be adapted to be used in an adaptive context if the automatic treatment planning is run on to generate new treatment plans on repeated images. In Chapter 5, CBCT images acquired at every treatment fraction and representing the patient anatomy of the day are used to compare the performance of three different automatic IMPT planning strategies based on dose mimicking. Although the three proposed strategies use dose mimicking for treatment plan optimization, the dose to mimic and the contours given as input to the optimizer are different. The evaluation of the strategies aims at investigating the benefits and limitations of the proposed methods for a potential use in an online adaptive setting.

Can Monte Carlo dose calculation be used in a fast online adaptive IMPT workflow?

Monte Carlo algorithms are the gold standard for accurate dose calculation in radiation therapy. The drawback of these methods is that they usually require very long computation times, which hampers their clinical application in an online workflow. Instead, most proposed online APT strategies use analytical algorithms to compute the dose. However, these algorithms may fail to predict proton range accurately in very heterogeneous media, such as lungs. Therefore, although we could benefit from a fast analytical dose calculation in the online setting (Chapter 2), the introduction of fast and accurate Monte Carlo dose calculation in commercial TPS opens the path for to explore this dose engine also for online re-optimization. The feasibility of OAPT strategies based on Monte Carlo dose calculation is described in Chapters 3 and 5.

Can online APT solutions be fully automated on a commercial TPS?

For an online adaptive workflow to be a candidate for clinical implementation, it needs to accomplish several requirements; it needs to be accurate, robust, fast, and automatic. The accuracy, robustness, speed, and automation of all the proposed strategies are discussed in this thesis. We aim to implement fully automated online APT workflows for two locations of greatest interest for adaptive IMPT: lung and HNC. The automation of the methods is performed via the scripting module of a commercial TPS, which includes robust optimization and Monte Carlo dose calculation. The full automation in a commercial TPS, well known by clinicians, could facilitate the clinical usage of the proposed solutions since proton treatment plan adaptation could be run with a single mouse click.

Chapter 2

Online dose restoration against density changes in lung IMPT

The content presented in this chapters correspond to the published article: [Online adaptive dose restoration in intensity modulated proton therapy of lung cancer to account for inter-fractional density changes](#). Physics and Imaging in Radiation Oncology (phiRO), 15:30–37, July 2020. Part of the special issue: Physics highlights from ESTRO 2020.

Authors: Elena Borderias-Villarroel¹, Xavier Geets ^(1,2) and Edmond Sterpin^(1,3)

¹ *UCLouvain, Institut de recherche expérimentale et clinique, Molecular Imaging and Radiation Oncology Laboratory, Brussels, Belgium.* ² *Cliniques Universitaires Saint Luc, Brussels, Belgium.* ³ *KULeuven, Department of Oncology, Laboratory of external radiotherapy, Leuven, Belgium.*

Summary

Background and purpose: In proton therapy, inter-fractional density changes can severely compromise the effective delivery of the planned dose. Such dose distortion effects can be accounted for by treatment plan adaptation, that requires considerable automation for widespread implementation in clinics. In this study, the clinical benefit of an automatic online adaptive strategy called dose restoration (DR) was investigated. Our objective was to assess to what extent DR could replace the need for a comprehensive offline adaptive strategy.

Methodology: The fully automatic and robust DR workflow was evaluated in a cohort of 14 lung IMPT patients that had a planning-CT and two repeated 4D-CTs (rCT1,rCT2). Initial plans were generated using 4D-robust optimization (including breathing-motion, setup and range errors). DR relied on isodose contours generated from the initial dose and associated patient specific weighted objectives to mimic this initial dose in repeated-CTs. These isodose contours, with their corresponding objectives, were used during re-optimization to compensate proton range distortions disregarding re-contouring. Robustness evaluations were performed for the initial, not-adapted and restored (adapted) plans.

Results and conclusions: The resulting DVH-bands showed overall improvement in DVH metrics and robustness levels for restored plans, with respect to not-adapted plans. According to CTV coverage criteria ($D_{95\%} > 95\% D_{\text{prescription}}$) in not-adapted plans, 35% (5/14) of the cases needed offline adaptation. After DR, Median($D_{95\%}$) was increased by 1.1 [IQR,0.4] Gy and only one patient out of 14 (7%) still needed offline adaptation because of important anatomical changes. DR has the potential to improve CTV coverage and reduce offline adaptation rate.

Highlights

- An online adaptive workflow based on automatic dose restoration (DR) was implemented.
- DR aims at compensating density changes by reproducing the planned dose.
- DR mitigates the impact of inter-fractional density changes in lung IMPT treatments .
- The need of comprehensive offline adaptation is reduced by introducing DR.
- DR improves target coverage and robustness respect to no adaptation.

2.1 Introduction

The sharp distal dose fall-off of the proton beams makes proton therapy (PT) an attractive alternative to conventional X-ray radiotherapy (XT). In particular, lung cancer patients may benefit from PT, as critical organs at risk (OAR) like the heart and oesophagus may surround the target. Studies showed that, compared to XT, intensity-modulated proton therapy (IMPT) has potential to improve OAR sparing whilst maintaining, or even increasing the prescribed dose to the target [204, 205, 206, 14, 24] and thus lead to better clinical outcomes i.e. less toxicity and higher survival rates [207, 208, 209, 210, 51].

The sharp dose gradients produced by IMPT yields dose distributions with increased sensitivity to treatment uncertainties such as patient setup, inter- and intra-fractional anatomical variations [15, 16]. This issue should not be addressed as typically done in XT, i.e. expansion of the clinical target volume (CTV) with safety margins leading to the planning target volume (PTV). The main reason is the invalidity in PT of the static dose cloud approximation underlying the PTV margin recipes. Therefore, more complex treatment planning procedures have been introduced such as robust optimization (RO) [97, 73, 72, 211, 212], where uncertainties are simulated during plan optimization. The interplay between respiratory and scanning motions must also be addressed cautiously, for instance using rescanning [213, 62].

Unfortunately, RO has several limitations. It is a computationally-intensive procedure, based on uncertainty parameters that are not patient specific (setup and range errors), which typically excludes important types of errors like day-to-day anatomical variations. However, those anatomical variations may have a critical impact, since about 30% of IMPT lung patients require adaptive planning even though they were initially planned using RO [24, 30]. Furthermore, some studies report that adaptation is highly recommended [212] or even mandatory [22] for advanced lung cancer treated with IMPT, since it ensures target coverage and improves treatment accuracy from a dosimetric point-of-view.

The degradation of plan quality observed in PT is caused by 1) dose distortions due to density changes along the beam path (PT specific) and 2) relative movement between anatomical structures. In order to handle both issues, a full adaptive workflow is necessary. However, its implementation requires considerable automation for a practical usage in clinical practice, typically characterized by tight clinical workflows. Automation is not straightforward, as it requires the generation of reliable new contours and a new plan, together with their associ-

ated verification and approval. The lack of automation is one of the key reasons preventing the deployment of online adaptive on a wide scale [155, 214].

As an alternative, dose distortions caused by density changes could be compensated by restoring the clinically approved planned dose. This approach, called 'dose restoration' (DR), was first introduced for online adaptation of prostate cancer [34], and later tested for head-and-neck and one lung tumors [33]. DR aims at reproducing the planned dose using density information from the new (repeated) CT, by generating a new (restored) plan with updated proton beam energies and spot weights, thereby removing the need for contour deformations. In such approach, automation was facilitated. Differences in dose distributions can indeed be quantified without qualitative evaluation by an expert, which is not the case for auto-segmentation. However, the clinical value and robustness of DR in lung remains to be assessed on a larger patient cohort. In this study, we evaluated the clinical value of the restoration workflow for 14 lung IMPT patients. Limitations with respect to a full adaptive workflow, where changes in relative positions and shapes of anatomical structures are considered, were also addressed.

2.2 Materials and methods

2.2.1 Lung patients data base

Fourteen lung cancer patients treated with photons were re-planned retrospectively using IMPT. A 4D-planning-CT and two repeated 4D-CTs (rCT1, rCT2) acquired systematically after 2 and 4 weeks of radiotherapy were available for each patient. For patient 9, rCT2 was missing. Patients were representative of a patient population with different tumor sizes (79–433 cm³) and motion amplitudes (1.1–11.1 mm), more characteristics can be found in the Appendix A, Table A.1. A joint ethical committee of Cliniques Universitaires St. Luc and UCLouvain approved the collection of the retrospective data used in this study.

2.2.2 Initial treatment planning strategy

Following the European Organization for Research and Treatment of Cancer (EORTC) protocol [215], a 5 mm expansion was applied to the GTV (Gross Tumor Volume) to create the CTV, and corrected to account for anatomical boundaries. The treatment planning system (TPS) RayStation 8A (RaySearch) was used to perform 4D worst-case robust optimization against breathing motion, setup and range errors. Robust parameters of 5 mm and 3% for setup and range errors, respectively, were used and extracted from literature [212, 216, 217]. Motion was modelled by using 3 images: the mid-position-CT (MidP-CT), used as planning

CT, together with the end-inhale and end-exhale breathing phases [218]. This combination of motion, setup and range errors led to 63 optimization scenarios : 7 (setup: $\pm 5\text{mm}$ in x,y,z directions plus nominal scenario) $\times$ 3 (density: $\pm 3\%, 0\%$) $\times$ 3 (three breathing-phases). The treatment plans were optimized using pencil beam dose engine (v4.2) and setting robust maximum and minimum objectives on the CTV. TPS calculations/optimizations were run on 2 Intel(R) Xeon(R) CPU-E5-2643-v3 3.40 GHz processors with 64 GB of RAM. A relative biological effectiveness (RBE) of 1.1 was assumed. The prescribed dose (D_p) to the CTV was 66 Gy (33×2 Gy).

2.2.3 Our proposal: the dose restoration (DR) workflow

Figure 2.1 illustrates the proposed adaptive workflow, including the online adaptive branch with automatic DR that did not require re-contouring on rCTs. The dose restoration approach used here was the “isodose volume dose restoration”, introduced by Bernatowicz et al. [33]. For preparing the DR, an additional step was performed during the planning stage compared to a regular planning procedure (see section 2.2.4). Online DR must be automatic, fast (patient is lying on the couch) and accurate (generates high quality plans that minimize the physician approval step). It is important to remember that DR only addressed distortions due to density changes. DR can account for tumor shrinkage for example, while for large tissue deformations and geometrical reorganization of the anatomy, an offline adaptation is necessary.

2.2.4 The isodose based optimized plan

First, a set of isodose volumes was constructed from the initial dose distribution obtained using the method described in section 2.2. Isodose volumes were created every 2 Gy within 95%–107% of the prescription dose and 5 Gy outside. Subsequently, an isodose-based optimization (IBO) was performed using a pair of minimum and maximum objectives for each isodose volume. Finally, the objective weights were manually adjusted in order to match IBO and initial plan DVHs. The objectives at isodoses corresponding to dose prescription were set as robust, using the initial robustness parameters (5 mm/3%/3phases). The resulting weighted objectives were saved in a patient-specific template, which served later to optimize the restored plan. Hence, each patient had two plans at the planning stage: the initial and the IBO plan. We assumed the IBO dose as the approved and delivered reference dose since the IBO and initial dose distributions were equivalent i.e. they both met clinical goals and presented overlapping DVHs.

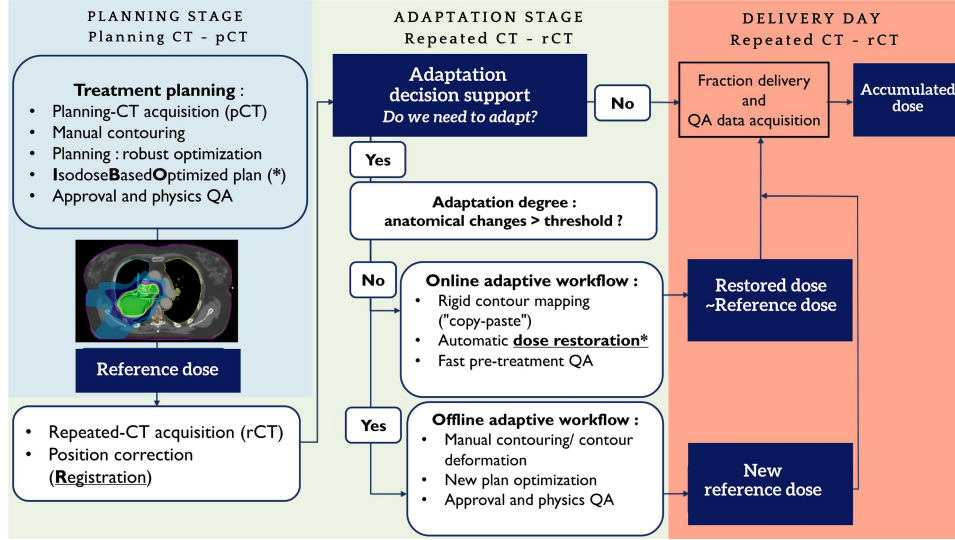


Figure 2.1: Example of the proposed adaptive workflow, which combines offline and online-adaptive strategies including dose restoration to perform optimal IMPT proton treatment. Three different stages are presented in this workflow: planning, adaptation and delivery day. If online adaptation strategy is chosen, the adaptation stage is performed on the delivery day. (*): The Isodose Based Optimized (IBO) plan is the key that automatizes dose restoration re-optimization.

2.2.5 DR on the treatment day: image registration, contour mapping and re-optimization

Bony-anatomy based rigid registration was performed between planning MidP-CTs and repeated MidP-CTs for patient positioning. Bony matching has been historically used in PT clinical routine to reproduce the relative position of bones with respect to protons [219, 220]. Using the registration matrix, contours (isodoses, OAR and target) were mapped from the planning MidP-CT to the repeated MidP-CTs. Afterwards, the previously calculated IBO template was used to re-optimize the plan.

2.2.6 Evaluation metrics

The impact of density changes was analysed by evaluating the treatment plans on two repeated 4D-CTs (rCT1, rCT2). The reference plans (IBO-plans), were recalculated on repeated-CTs. Not adapted dose distributions were compared to restored doses using dose-volume parameters such as D95%(CTV) (dose received by 95% of the target volume), and the homogeneity index $HI = (D95\% - D5\%) / D50\%$. Clinical limit for target coverage was defined at 95% of the prescribed dose, i.e. $D95\%(CTV) > 95\%D_p (66 \text{ Gy}) = 62.7 \text{ Gy}$.

For robustness evaluation, the reference (IBO), restored and not adapted plans were re-computed in the same 63 scenarios used for optimization (5 mm, 3%, 3 breathing phases) using RayStation dose perturbation tool. The DVHs from the 63 simulated uncertainty scenarios were grouped into DVH-bands. The band-width served to measure the robustness level. Propagation of the robustness level through DR was analysed. For each DVH-metric (D95%(CTV), D2%(SpinalCord) or mean dose in OAR) nominal and worst-case scenarios values were reported. Differences respect to the reference dose were compared for not adapted and restored plans.

Dose-volume parameters and robustness were evaluated in the rigidly mapped contours (RS_{map}) and in the contours (OAR and targets) recontoured by an experienced radiation oncologist. This last set will from this point onward be denoted as 'real' structures (RS_{real}). Evaluations were performed for the two repeated-CTs (rCT1, rCT2).

Dice similarity coefficient (DSC) was calculated between CTV_{map} and CTV_{real} to quantify the accuracy of the rigid registration between planning and repeated-CTs (Appendix A, Table A.2). This statistical tool which measures the similarity between two data sets X, Y ($2|X \cap Y|/(|X| + |Y|)$) is often used to evaluate the accuracy of segmentation methods.

To ensure a reliable and accurate dose mimicking, we also analysed local differences respect to the reference (IBO) dose in four different dose level regions (prescribed, high, medium and low dose) following Bernatowicz et al.(2018)[33]. Subtractions between the evaluated (not adapted/adapted) and reference doses were performed. Absolute dose difference at 2% of the volume was reported for each region (for instance $DE(vol = 2\%) = 5$ Gy means that only 2% of the volume may have a dose difference superior to 5 Gy). This metric characterized the near maximum absolute dose error.

Analysis were performed using boxplots calculated with values of the 14 patients. Medians [interquartile range(IQR)] from not adapted versus restored doses in rCT1 and rCT2 were reported and compared using Wilcoxon signed-rank test (paired difference test), in order to assess the statistical significance. Differences between not adapted and restored metrics were considered statistically significant for a p-value < 0.05 .

2.3 Results

2.3.1 DR improves target coverage and homogeneity index

Figure 2.2 presents $D95\%(CTV)$ and homogeneity index (HI) values evaluated in the nominal scenario, which represents the scenario without added uncertainty. Even in this case, 35% (5/14) in rCT1 and 23% (3/13) in rCT2 of patients would have benefit from adaptation to ensure clinical acceptable target coverage.

As observed in Figure 2.2, reference plans were significantly distorted without adaptation, leading to CTV underdosage and higher HI in both repeated-CTs. Both metrics were improved after restoration, showing that DR can deal with over/under-dosages.

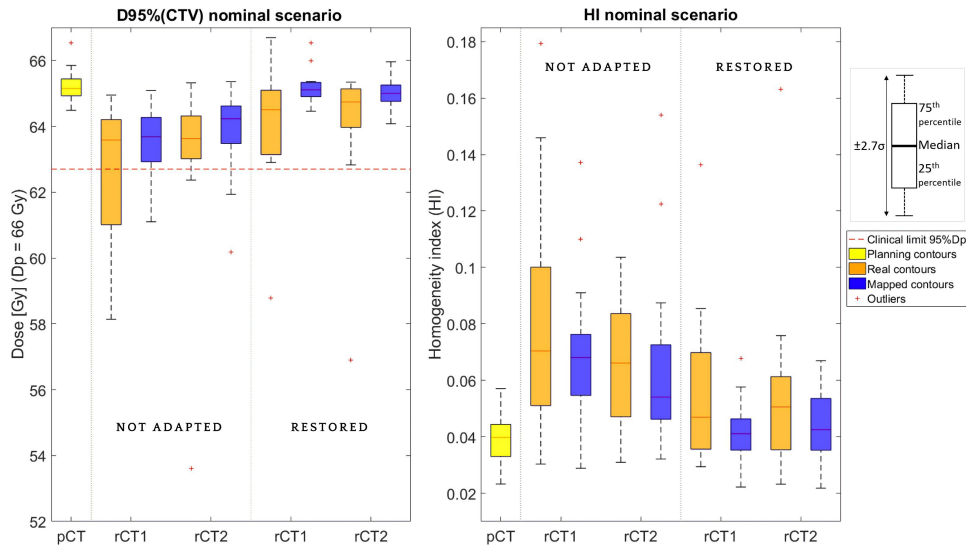


Figure 2.2: $D95\%(CTV)$ (dose received at 95% of the CTV) and HI values from nominal scenario are represented using boxplots. The nominal scenario stands for doses without any simulated uncertainty, which means without robust assessment. Each boxplot contains the 14 patient's information for a certain plan (not adapted/restored) evaluated on different images: planning CT and two repeated-CTs. The clinical limit is achieved if dose received at $D95\%(CTV)$ is higher than $95\%(Dp(66 \text{ Gy})) = 62.7 \text{ Gy}$. Abbreviations: Dp =Prescribed dose, HI =Homogeneity index, pCT =planning CT, $rCT1$ =first repeated CT, $rCT2$ =second repeated CT.

Without adaptation, in RS_{map} , $\text{Median}(D95\%(CTV))$ was reduced from 65.2[0.5] Gy to 63.7[1.39] Gy (rCT1) and to 64.2[1.28] Gy (rCT2). For the RS_{real} , $\text{Median}(D95\%(CTV))$ fell to 63.6[3.4] Gy (rCT1) and 63.6 [1.4] Gy (rCT2). After restoration, target coverage was also re-established increasing $\text{Median}(D95\%(CTV))$

to 65.1[0.4] Gy (rCT1) and to 65.0[0.5] Gy (rCT2) in RS_{map} and to 64.5 [2.0] Gy (rCT1) and to 64.7[1.3] Gy in RS_{real} . In all restored cases (RS_{map} and RS_{real}), D95%(CTV) values were maintained within clinical limits, except for RS_{real} in patient 6. In this case, the rigid registration failed due to substantial differences in the anatomy. As a result, the RS_{real} did not match RS_{map} (DSC of 0.36 (rCT1) and 0.32 (rCT2)), and the initial dose was reproduced in the wrong location. Only in this case, offline adaptation was absolutely needed to ensure CTV coverage.

2.3.2 DR increased also robustness levels

Figure 2.3 shows the deviations from the reference dose for not adapted and restored plans in nominal and worst-case scenarios for D95% (CTV), D2% (Spinal-Cord) and MeanDose (Lung/Heart/Esophagus).

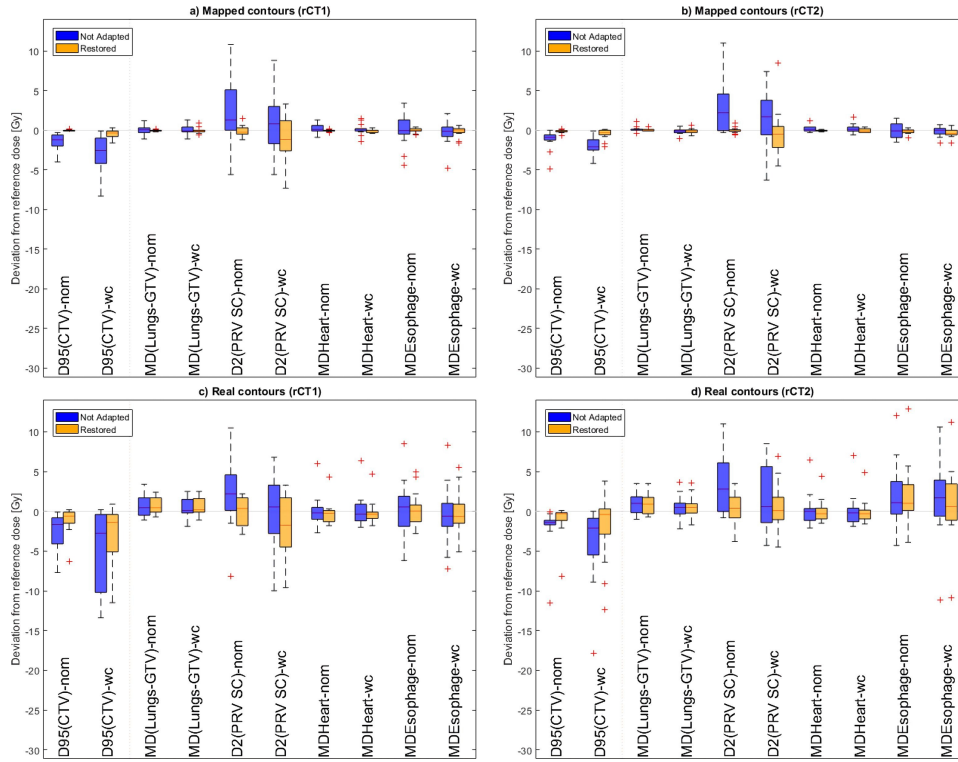


Figure 2.3: Differences respect to the reference dose in nominal and worst-case metrics reported from DVH-bands were represented using boxplots. The median calculated among fourteen patients is shown by the horizontal line within boxes. Results were evaluated for both repeated-CTs(rCT1,rCT2) in: mapped contours (RS_{map}) a),b) and real contours (RS_{real}) c),d). Abbreviations: nom = nominal, wc = worst case, MD = mean dose, D2 = D2% dose received at 2% of the volume; PRV SC = Spinal Cord.

After adaptation, median values were closer to zero, indicating that the reproduction of the initial dose was improved. Some OAR (e.g. spinal cord) even got lower doses. Adapted plans also showed improved results in worst-case DVH-metrics.

Larger variabilities were observed in RS_{real} (Figure 2.3.c and .d) compared to RS_{map} (Figure 2.3.a and .b). However, in both contour sets, all restorations improved target coverage and recovered OARs metrics closer to original values. The number of outliers increased in RS_{real} , with most outliers corresponding to patient 6, who had significant CTV underdosage and heart overdosage.

Outliers representing an overdosage in the esophagus on rCT2 (Figure 2.3.d) were associated to the entrance of the esophagus into a high dose region, caused by tumor shrinkage. Underdosage in the CTV and overdosage in D2(PRV SC) worst-case in rCT2 (Figure 2.3.d) correspond to patient 8. In this case, the robustness level was not achieved in rCT2, being translated into a wider CTV DVH-band. This is probably due to contour miss-alignment characterized with a DSC of 0.67.

There was no significant dose difference for lung, heart and esophagus between not adapted and restored plans ($p > 0.05$). However statistical significance ($p < 0.05$) was observed in target coverage and maximum dose (D2%) in the spinal cord, for nominal and worst-case values (see Table 2.1).

	Mapped contours (RS_{map})		Real contours (RS_{real})	
	p (rCT1)	p (rCT2)	p (rCT1)	p (rCT2)
D95(CTV)-nom	0.0001	0.0010	0.0002	0.0005
D95(CTV)-wc	0.0005	0.0024	0.0295	0.0029
MD(Lung-GTV)-nom	0.6533	0.9063	0.8325	0.6914
MD(Lung-GTV)-wc	0.9679	0.6328	0.9883	0.8926
D2(SpinalCord)-nom	0.0166	0.0134	0.0580	0.0269
D2(SpinalCord)-wc	0.1465	0.2163	0.0906	0.2930
MD(Heart)-nom	0.1328	0.0811	0.2251	0.3955
MD(Heart)-wc	0.1851	0.1230	0.2578	0.2490
MD(Esophagus)-nom	0.9863	0.6563	0.7494	0.7764
MD(Esophagus)-wc	0.9604	0.3696	0.8394	0.5049

Table 2.1: Summary of Wilcoxon's signed rank test p-values in the two contours sets : mapped contours (RS_{map}) and real contours (RS_{real}) for both repeated-CTs(rCT1, rCT2). *Abbreviations* : D95(CTV) = dose received at 95% of the CTV, MD = mean dose, D2(SpinalCord) = dose received at 2% of the volume of the spinal cord, nom = nominal, rCT = repeated CT, wc = worst-case.

Figure 2.4 shows an example (Patient 9) of CTV underdosage close to clinical limit, causing the robustness test to fail. After DR, target coverage was improved and robustness level was recovered close to initial values.

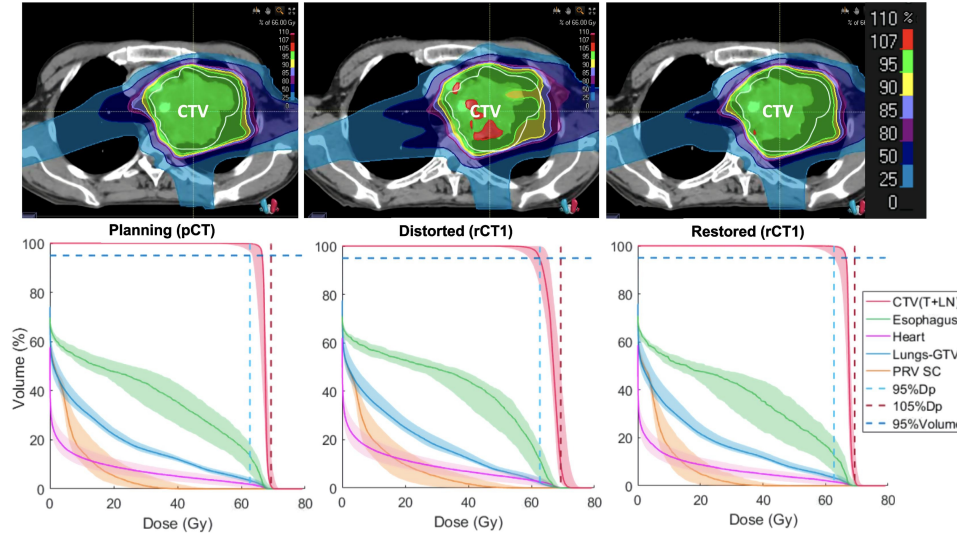


Figure 2.4: Patient 9. DVH-bands resulting from the robustness tests are presented together with their corresponding nominal dose distributions. The continuous line represents the nominal scenario while the band collects the information from the evaluated uncertainty scenarios. The dashed lines represent the clinical limits ($D_{95\%} > 95\%D_{prescribed}$). The pair of results (dose distribution/DVH-bands) are shown for the initial IBO plan (in the planning CT), the not adapted and the restored plans (in the first repeated-CT). Abbreviations: *pCT* = *planning-CT*, *rCT1* = *first repeated-CT*, *rCT2* = *second repeated-CT*; *CTV(T + LN)* = *CTV(tumor and nodes)*; *PRV SC* = *Spinal Cord*; *Dp* = *Prescribed dose*.

2.3.3 Local dose differences were reduced using DR

Figure 2.5 presents near maximum dose errors ($DE_{vol} = 2\%$) in four different regions: prescription ($>95\%D_p$), high ($65-95\%D_p$), medium ($35-65\%D_p$) and low ($<35\%D_p$) dose levels. DR could reproduce the dose with high accuracy and precision in the whole patient volume, with $\text{Median}(DE_{vol} = 2\%) < 4$ Gy in all dose level volumes. Dose differences between not adapted and restored doses were statistically significant ($p < 0.05$) in all regions for both rCTs. $\text{Median}(DE_{vol} = 2\%)$ of the entire dose region was decreased from 9.2[5.2] Gy (rCT1) and 8.2[4.7] Gy (rCT2) in the not adapted plans to 3.1[2.3] Gy (rCT1) and 2.9[1.7] Gy (rCT2) in the restored plans. Additionally, DR precision, represented by the IQR, was decreased by a factor two respect to not adapted plans.

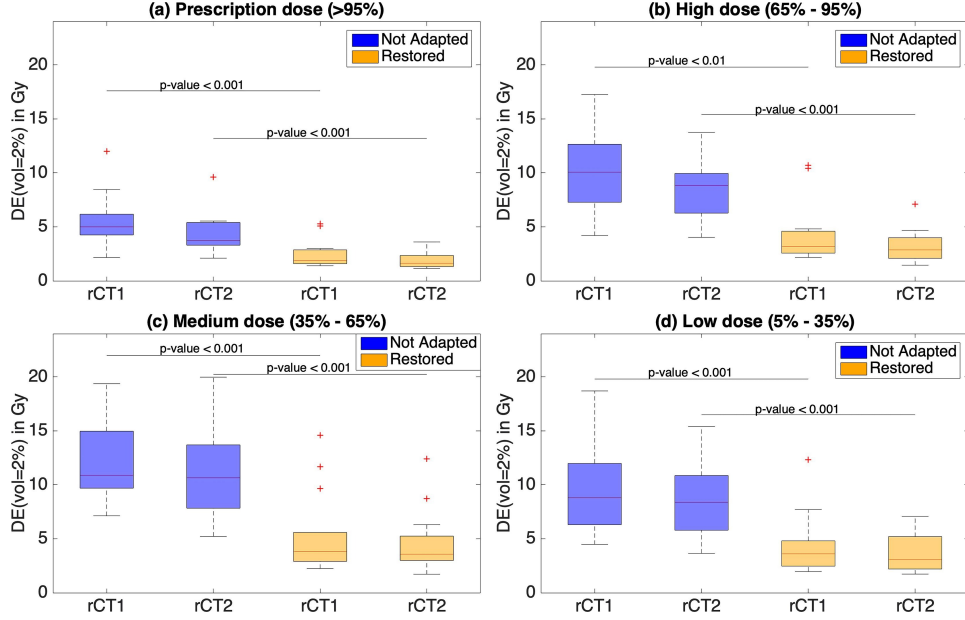


Figure 2.5: Absolute dose error respect to the reference dose in small volumes $DE(vol = 2\%)$ represented in boxplots for 4 different regions corresponding to prescription, high, medium and low dose levels. Abbreviations: *rCT1* = first repeated-CT, *rCT2* = second repeated-CT.

2.4 Discussion

In this study, the clinical value of the DR tool for lung IMPT accounting for inter-fractional density changes, in order to reduce the need of offline adaptation was evaluated. We observed that introducing DR reduced offline adaptation rate from 36%(5/14) to 7%(1/14).

Fast and accurate DR will be sufficient for the majority of the clinical cases. However, issues such as (1) changes in the relative positions of anatomic structures or (2) poor quality registrations may cause an imperfect overlap between rigidly mapped and real contours. In such cases, the positive effects of DR will be less noticeable despite a successful reproduction of the initial dose, since it will be reconstructed in the wrong location. The first issue is not unique to PT, it could be solved as in XT, using a more comprehensive offline adaptation including re-contouring. For online adaptation, extensive quality assurance (QA) of tools as auto-contouring [183, 221] or automated planning techniques [222, 223] would be needed. A possible solution for (2) could be other registration strategies such as soft-tissue or carina-based positioning [224]. Bony registration could be preferred in current practice for ensuring correct alignment of the bones with respect to

proton paths and avoid dose distortions. However, soft tissue or carina based positioning could be implemented safely due to inherent corrections brought by restoration.

Based on local dose differences analysis, regions of prescribed dose consistently showed better performance with the lowest $\text{Median}(\text{DEvol} = 2\%) = 1.7 \text{ Gy}$ in restored plans. These results, together with the $\text{Median}(\text{DE}(2\%\text{vol})) < 4 \text{ Gy}$ for entire dose region, were consistent with $\text{Median}(\text{DE}(2\%\text{vol})) < 5 \text{ Gy}$ obtained by Bernatowicz et al. [33].

Uncertainty sources such as motion, setup and range errors were included in the DR robust re-optimization. DR could improve nominal and worst-case DVH-metrics in RS_{map} as well as in RS_{real} respect to no adaptation. Lower differences in worst-case values from the DVH-bands of restored plans indicated that adapted plans were more robust than not adapted plans.

Interplay effect was not addressed in this study since it is not a problem specific to DR but to IMPT for mobile-targets in general. Issues associated to interplay have been studied and well known solutions such as re-scanning have been proposed in the literature to compensate this effect [217, 225, 63, 62].

As investigated in Van Der Voort et al. [211] for oropharyngeal cancer, the magnitude of range and setup robustness parameters can be approximated as a function of systematic and random uncertainties. Hence, by increasing the initial plan robustness, less adaptation should be required. The implementation of an online adaptive workflow with DR opens the possibility to reduce margins and/or decrease robustness parameters, compared to non-adaptive workflows. However, the relationship between the robustness parameters and the adaptation rate must be studied carefully on a suitable patient cohort.

Using DR, the initial plan was adapted to the new patient density. The original plan was preserved (beam arrangement, target volumes) whilst re-optimization allows for modifications in proton energies and spot weights in order to reproduce the original dose. Keeping the reference dose distribution unchanged is the key to minimize the physician's approval step in a fully online adaptive workflow. However, despite restoring the approved planned dose, online adaptation requires modification of the pre-treatment patient-specific QA procedure. For instance, a previous QA could be performed using a fast-independent dose calculation, followed by a more extensive analysis using machine log files after delivery [104, 102, 151]. Both procedures (pre/post treatment) could identify patients that require offline adaptation because of anatomical changes DR cannot account for.

Clinical implementation of DR, will mostly impact the workload at the planning stage since the IBO plan needs to be optimized. Pencil Beam dose engine was used to speed up the workflow since Monte Carlo optimization was not needed to prove the suitability of the method, which focused more on the clinical feasibility of automating the process than on speed. DR workflow was able to generate plans of acceptable quality (accounting for density changes), in around 30 min [28 min 20 s – 31 min 21 s]. Since re-optimization is the most time demanding step, advanced optimization techniques, as scenario selection [98] and added computational power can make the time scale compatible with typical online treatment settings.

By routinely acquiring high quality patient volumetric images (e.g. in-room CT on rails), systematic setup errors can be mostly mitigated, allowing DR to deal primarily with the range errors. However, on-board CBCT could also be considered, but this requires advanced image quantification features in order to use CBCT information for dose computation [142, 226, 143].

Dose differences between not adapted and restored plans were statistically significant ($p < 0.05$) in all dose regions, the target volume and the spinal cord. No statistical signification was found for heart, lung and esophagus. However, this is not surprising since dose deviations between not adapted and planned dose distributions were in general small in organs-at-risk. In addition, it must be noted that patients who did not need adaptation were also included in our statistical analysis. More statistically significant results could have been achieved if we had included only patients that needed adaptations. Yet, we achieve statistical significance for target volumes, all dose regions, and the spinal cord.

Finally, performing fast and accurate DR improved CTV coverage for all patients in comparison with no adaptation. DR key objective was to meet the planning criteria set by the radiation oncologists which were judged as clinically relevant. Since restoration allowed to get closer to the validated initial plan, and sometimes reestablish the planning criteria, this improvement is, by definition, clinically relevant. Ideally, DR could be considered as standard clinical practice, in the same way as daily registration based on volumetric images.

In conclusion, most of the investigated lung cases would benefit from DR. This online adaptive workflow (based on rCTs) did not require new contour segmentation nor deformation. DR was validated in rigidly mapped contours as well as in the real delineated contours by the physician. DVH-metrics and robustness were improved by DR in the new patient anatomy compared to no adaptation.

Chapter 3

Clinical value of online dose restoration for head and neck cancer

The content presented here comes from the published article: [Evaluation of the clinical value of automatic online dose restoration for adaptive proton therapy of head and neck cancer](#). Radiotherapy and Oncology, 170:190–197, 2022.

The publication of this paper was the result of a collaboration with the MAASTRO Clinic. The stay at this institution started onsite for a short period of one month but had to continue online due to health restrictions.

Authors: Elena Borderias-Villarroel¹, Vicki Taasti², Wouter Van Elmpt², Sara Teruel-Rivas¹, Xavier Geets^(1,3) and Edmond Sterpin^(1,4)

¹ UCLouvain, Institut de recherche expérimentale et clinique, Molecular Imaging and Radiation Oncology Laboratory, Brussels, Belgium. ² Department of Radiation Oncology (MAASTRO), GROW – School for Oncology, Maastricht University Medical Centre, Netherlands. ³ Cliniques Universitaires Saint Luc, Brussels, Belgium. ⁴ KULeuven, Department of Oncology, Laboratory of external radiotherapy, Leuven, Belgium.

Summary

Background and purpose: Intensity modulated proton therapy (IMPT) is highly sensitive to anatomical variations which can cause inadequate target coverage during treatment. This study compares not-adapted (NA) robust plans to two adaptive IMPT methods - a fully-offline adaptive (FOA) and a simplified automatic online adaptive strategy (dose restoration (DR)) to determine the benefit of DR, in head and neck cancer (HNC).

Methodology: Robustly optimized clinical IMPT doses in planning-CTs (pCTs) were available for a cohort of 10 HNC patients. During robust re-optimization, DR used isodose contours, generated from the clinical dose on pCTs, and patient specific objectives to reproduce the clinical dose in every repeated-CT(rCT). For each rCT($n = 50$), NA, DR and FOA plans were robustly evaluated.

Results: An improvement in DVH-metrics and robustness was seen for DR and FOA plans compared to NA plans. For NA plans, 74%(37/50) of rCTs did not fulfill the CTV coverage criteria ($D_{98\%} > 95\% D_{\text{prescription}}$). DR improved target coverage, target homogeneity and variability on critical risk organs such as the spinal cord. After DR, 52%(26/50) of rCTs met all clinical goals. Because of large anatomical changes and/or inaccurate patient repositioning, 48%(24/50) of rCTs still needed full offline adaptation to ensure an optimal treatment since dose restoration was not able to re-establish the initial plan quality.

Conclusion: Robust optimization together with fully-automatized DR avoided offline adaptation in 52% of the cases. Implementation of dose restoration in clinical routine could ensure treatment plan optimality while saving valuable human and material resources to radiotherapy departments.

Highlights

- An online adaptive strategy was compared to not-adaptation of robust clinical plans and offline adaptation.
- Not-adapted robustly optimized plans achieved adequate target coverage in up to only 26% of repeated-CTs.
- Dose restoration yielded adequate target coverage in 52% of repeated-CTs
- Dose restoration reduced the dose to most organs-at-risk compared to not-adaptation of robust clinical plans.
- Full offline adaptation was still required in 48% of repeated-CTs.

3.1 Introduction

Intensity-modulated proton therapy (IMPT) has shown the potential to reduce the dose to the organs at risk (OAR), while keeping a similar target coverage compared to intensity-modulated radiation therapy (IMRT) [227, 228] and volumetric modulated arc therapy (VMAT) [228, 229] in head and neck cancer (HNC). Lower dose levels in critical OAR translates into reduced radio-induced toxicities such as xerostomia, dysphagia, or tube feeding dependence [230, 231, 50, 232, 233]. However, IMPT is highly sensitive to patient anatomical changes due to the nature of proton interaction with tissues. The initially planned Bragg Peak positions can be altered by density changes along the beam path caused by tumor shrinkage, weight loss, positioning errors, or shifts in tumor and organ locations. Fraction after fraction, these uncertainties can cause distortions on the approved clinical plan and compromise plan quality [15, 16, 17].

Different techniques have been proposed to mitigate the distortion of the initial plans due to inter-fractional density changes and anatomical deformations: robust treatment planning, offline and online adaptive IMPT. Robust optimization (RO) strategies include uncertainties during spot weight optimization and lead to dose distributions that are resilient to specific uncertainties, classically systematic setup and range errors [77]. More sophisticated approaches, such as anatomical RO, also include anatomical scenarios during plan optimization [124, 234, 125]. Nevertheless, it remains difficult not to compromise the optimality of the IMPT plan while trying to cover a wide range of possible anatomical changes. RO can provide robustness against some density and anatomical changes but it is not the ultimate solution, treatment plan adaptation is still required [229, 17, 200, 235].

A typical full offline adaptation can take several days in the clinical workflow since it requires re-contouring, re-planning, treatment plan approval and additional quality assurance. In the meanwhile new anatomical changes could occur. In order not to delay treatment delivery, another possibility is to perform an online plan adaptation while the patient is on the treatment couch. Implementation of full online adaptation is very challenging since it requires completing the adaptation workflow steps in a very short time. Ideally, re-contouring would rely on automatic contour propagation based on artificial intelligence tools and/or deformable image registrations but for the moment they cannot be used independently without human intervention [236], and would therefore still require approval by the doctor.

As an alternative, online adaptive strategies based on the restoration of the dose have been proposed [33, 34]. The idea is to keep the planned dose distribution stable in every fraction by generating a new (restored) plan with updated proton

beam energies and spot weights. The advantage of these strategies is that they do not require new contour delineation prior to re-optimization, contours from the planning-CT are instead rigidly mapped. Promising results illustrate the potential to mitigate inter-fractional dose distortions while maintaining target coverage for lung [237] and prostate cancer [238].

This study compares a not adapted robust planning strategy (RO) with two active adaptive IMPT strategies: a fully-offline adaptive (FOA) strategy and a robust automatized online adaptive strategy based on dose restoration (DR). The added clinical value and/or limitations of dose restoration in HNC patients with respect to a full adaptive workflow, including the impact of changes in relative positions and shapes of anatomical structures, are investigated.

3.2 Materials and methods

3.2.1 Patient data and dose prescription

This study included data of ten HNC patients (see Table 3.1), with a planning CT scan (pCT) and four to six repeated CT scans (rCTs) per patient. Clinical treatment plans were generated at the Department of Radiation Oncology of Maastricht University Medical Center (MAASTRO Clinic, Maastricht, The Netherlands) and delivered using the MEVION S250i Proton Therapy System with HYPERSCAN Pencil Beam Scanning [239]. Institutional review board approval was granted for this study. The patients were treated with a prescription dose of 70 GyRBE to the high-dose clinical target volume (CTV) and 54.25 GyRBE to the low-dose CTV in 35 fractions, using a constant relative biological effectiveness (RBE) factor of 1.1. The spinal cord, parotids, submandibulars and the oral cavity were defined as OAR. Target and OAR delineations performed by experienced radiotherapy oncologists were available on all pCTs and rCTs.

3.2.2 Treatment planning strategies

Three different planning strategies aiming at clinically acceptable doses in each rCT were compared. Robust optimization (RO) planning without treatment adaptation and two active adaptation strategies: online dose restoration (DR) and full offline adaptation (FOA), were applied and evaluated as described in Figure 3.1. All dose calculations were run in RayStation 10A (Raysearch Laboratories, Stockholm, Sweden) using the Monte Carlo algorithm in a $3.0 \times 3.0 \times 3.0 \text{ mm}^3$ dose grid.

Patient	Adaptation	Location	N_beams	Days after pCT for each rCT					
				1rCT	2rCT	3rCT	4rCT	5rCT	6rCT
1	Y	Nasopharynx	7	20	27	34	41	48	55
2	Y	Larynx	4	25	33	39	47	-	-
3	Y	Oropharynx	5	17	23	30	36	44	-
4	N*	Oropharynx	5	16	23	30	37	44	-
5	Y	Oropharynx	5	15	22	28	35	41	44
6	Y	Oropharynx	4	24	29	38	44	50	-
7	Y	Hypopharynx	4	26	34	40	43	-	-
8	N*	Nasopharynx	9+	21	28	35	42	49	-
9	Y	Hypopharynx	5	18	22	29	36	43	54
10	Y	Oropharynx	5	10	17	24	31	-	-

Table 3.1: Characteristics of the head and neck IMPT patient data-base. Abbreviations: Y = Yes, N = No, pCT = planning-CT, rCT = repeated-CT, *Patients not being adapted during treatment but with clinical goals not achieved in some of the fractions. +Due to the limited field size two different isocenters were needed for this patient.

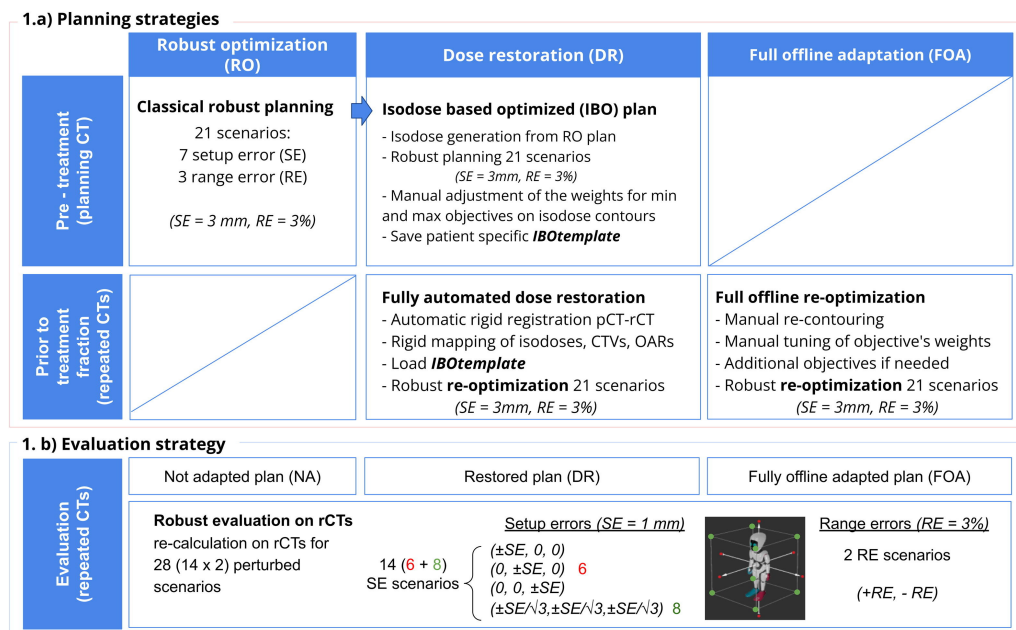


Figure 3.1: Summary of 1.a) the three planning strategies compared in this study and 1.b) the robustness evaluation.

3.2.3 Method 1: Robust treatment planning

A robustly optimized (RO) treatment plan was created on the pCT applying classical worst-case robust optimization (setup error (SE) = 3 mm, range error (RE) = 3%). SE scenarios included patient shifts in the three axes of the space x,y,z ($\pm SE, 0, 0$), $(0, \pm SE, 0)$, $(0, 0, \pm SE)$ and the nominal scenario $(0,0,0)$. Range robustness was also included to account for uncertainties in the conversion from CT numbers to proton stopping power ratios. RE scenarios were modeled by scaling the mass density by $\pm 3\%$. The nominal scenario (i.e. 0%) was also included. The combination of seven SE scenarios and three RE scenarios lead to an optimization with twenty-one (7×3) error scenarios.

Robust optimization was followed by a robust evaluation using the same robustness parameters (SE = 3 mm, RE = 3%). RO plans satisfied, in nominal and worst-case scenarios, the following clinical goals for both the high- and low-dose target volumes: D98% 95% and D2% 107% of the prescription dose, where D98% and D2% are the doses received at 98% and 2% of the target volume, respectively. For the OAR, the constraints were: $D_{0.03 \text{ cm}^3} < 45 \text{ Gy}$ in the spinal cord, $D_{\text{mean}} < 26 \text{ Gy}$ to both parotids, both submandibular glands and the oral cavity.

3.2.4 Method 2: Dose restoration (DR)

The dose restoration approach used in this study was the isodose volume dose restoration, an in-house developed workflow introduced by Bernatowicz et al. [33] and further investigated by Borderias et al. [237]. Dose restoration was fully-automatized using RayStation scripting but required the generation of an additional pre-treatment plan compared to a regular planning procedure (see Figure 3.1).

First, isodose contours were constructed from the clinical dose in the original plan on the pCT (every 2 Gy within 95%–107% of the prescription dose and 5 Gy outside). Later, still at the planning stage, optimization parameters (i.e., a set of weighted minimum and maximum objectives) were manually tuned in order to generate the so-called isodose-based optimized (IBO) plan, which reproduces the planned isodose distributions. Robust objectives were set for the isodoses corresponding to prescription dose levels (70 and 54.25 Gy) and the CTVs. The isodose levels and optimization parameters of the IBO plan were saved in a patient-specific template (IBO template) to be further used for fully-automatized dose restoration on rCT as follows.

After the acquisition of the rCT, automatic rigid registration based on the bony anatomy was performed between the pCT and the rCTs. Then, the previously generated isodose contours and the delineated CTVs and OAR on the pCT were rigidly mapped to the rCT. The clinical plan on the pCT was used as a warm-start for adaptation. The same beam configuration, beam computation settings, robustness parameters and spot position initialization were used for re-optimization together with the patient-specific IBO template generated on the pCT. Restored plans were characterized by updated spot weights and energies.

3.2.5 Method 3: Full offline adaptation

Instead of using rigidly mapped contours, FOA plans were manually and robustly re-optimized based on manually delineated contours available on the rCTs. The same optimization objectives as in the clinical plans on the pCT were used. Objective weights were manually adjusted to meet the clinical goals and extra optimization objectives were added if necessary.

3.2.6 Robust evaluation of the methods

The clinical plan, which is robustly optimized on the pCT, the DR plan, and the FOA plan were robustly evaluated on the rCTs. Robust evaluation was performed simulating 28 uncertainty scenarios (14 SE scenarios $\times$ 2 RE scenarios) for each of the plans. As described in Figure 3.1.b, the sampling of the setup error scenarios included the 6 principal directions (6 red dots) and the diagonals from the center to the vertices (8 green dots) as suggested in Korevaar et al. [101]. From this point on, we will refer to the evaluation (i.e. re-calculation) of the clinical plan on rCTs as the non-adapted (NA) plan, as this plan was not actively adapted. Plan parameters such as energy layers, number of spots or spot weights were not modified when recomputing the dose for this plan on the rCTs.

3.2.7 Evaluation metrics

D98% and D2% were the dose-volume parameters evaluated for both the high- and low-dose target volumes, while D0.03 cm³ was used for the spinal cord and mean doses for the other OAR. Since our main objective was to assess the reproduction of the clinical dose on the rCTs, we did not focus on absolute dosimetric values. For each metric, we instead analyzed the differences of the NA, DR and FOA plans with respect to the clinical dose for the nominal and the worst-case scenarios.

The nominal and worst-case values of the clinical dose were extracted from the robustness evaluation (SE = 3 mm, RE = 3%) of the RO plan on the pCT. Nominal (SE = 0 mm, RE = 0%) and worst-case values (SE = 1 mm, RE = 3%) out of

all scenarios of NA, DR and FOA plans were reported on the manually delineated contours. Median values and interquartile range (IQR) of the differences with respect to the clinical dose were calculated for each metric over the 50 rCTs.

Wilcoxon signed-rank tests were performed to assess the statistical significance of the differences between dose metrics for each strategy. We were interested in the performance of adaptation workflows (DR and FOA), especially when the not-adapted plans were distorted (i.e., not meeting the clinical goals on the rCTs). Hence, for cases where the not-adapted plan differed less than 0.5 Gy with respect to the clinical plan on the pCT were not included in the significance tests. A p -value < 0.05 was considered statistically significant.

The Dice similarity coefficient (DSC) was calculated between the manually delineated and rigidly mapped contours on the rCTs to quantify the accuracy of the rigid registration between the pCT and the rCTs.

3.3 Results

3.3.1 Analysis of DVH-metrics

For the NA plans, the clinical goals for the target coverage (i.e., D98% 95%) in the nominal scenario were only met in 26% (13/50) of the rCTs. After performing DR, the percentage was increased to 52% (26/50). All plans met all clinical goals when full off-line adaptation was applied.

Figure 3.2 shows the differences in DVH-metrics between the approved clinical plan on the pCT and the NA, DR, and FOA plans in the OAR and CTVs with two prescription levels: high-dose CTV and low-dose CTV. Target coverage (D98%) of the high-dose CTV was significantly lower for NA plans compared to the DR and FOA plans (both $p < 0.0001$, Appendix B.1, Table B.1), with median values of the difference with respect to the clinical plan of 0.8 (IQR = 1.3) Gy, 0.1 (IQR = 1.0) Gy and 0.1 (IQR = 0.2) Gy, respectively.

The over-dosage of the high-dose CTV, characterized by D2%, showed a similar trend, with median values of the difference with respect to the clinical plan of 0.53 (IQR = 1.34) Gy, 0.19 (IQR = 0.56) Gy and 0.03 (IQR = 0.36) Gy for the NA, DR and FOA plans, respectively. Once again, differences between DR and FOA with respect to NA were significant (both $p < 0.0001$, Appendix B.1, Table B.1).

Target coverage (D98%) of the low-dose CTV was lower than for the clinical plan for both NA and DR plans with the same median values but with a lower variability for the DR plans: differences of 0.7 (IQR = 2.3) Gy for NA and 0.7

(IQR = 1.2) Gy for DR. Analyzing only distortions greater than 0.5 Gy in NA plans, significant differences were found for DR plans compared to NA plans in nominal scenarios ($p = 2.1 \cdot 10^{-4}$). Differences in the worst-case D98% values between NA and DR were also significant for both CTVs ($p = 7.8 \cdot 10^{-4}$ and $p = 3.8 \cdot 10^{-4}$ for low-risk and high-risk CTVs, respectively).

Regarding critical OAR, such as the spinal cord, the DR plans offered a significant sparing, in comparison to the NA plans, for both the nominal scenario and worst-case scenario. Dose differences in the right submandibular gland and the oral cavity were not significant between the three methods in nominal or worst-case values.

3.3.2 DR can compensate for dose distortions

Figure 3.3 illustrates three patient examples where DR corrected the distortions in NA plans. Patient 3, who experienced a reduction of the lateral neck diameter, and patient 8, with a suboptimal re-positioning of the shoulders, presented significant under-dosages for the low-dose and high-dose CTV in the NA plan, respectively. DR was able to recover isodose 95% of the prescribed doses in both cases, improving dose conformality and robustness.

3.3.3 Limitations of DR for non-rigid deformations

However, outliers, showing critical under-dosage in CTVs (D98%) or increased dose in the spinal cord D0.03 cm³, highlighted the limitations of DR to improve plan quality in some cases. Dose distributions associated with two of these patients are shown in Figure 3.4.

For patient 1, cold spots were visible in the low-dose CTV of both the NA and DR plans in the second rCT. In this case, DR did reproduce the dose distribution on the rCT, but not at the correct location since the rigid registration between pCT and rCT could not compensate for the non-rigid deformations on the lower part of low-dose CTV due to an inaccurate shoulder repositioning (Appendix B, Figure B.1). This led to a sub-optimal overlap between the delineated and rigidly mapped low-dose CTV characterized by a DSC coefficient of 0.78. As shown for patient 2 (Figure 3.4.b), anatomical deformations involving relative movement between targets and OAR are also a cause of poor alignment between manual and rigidly mapped contours. High-dose CTV shrinkage of 11.5% together with a shift in the cranio-caudal direction and a more pronounced neck curvature in the rCT led to the spinal cord moving into the dose distribution (the DSC score between the manually delineated and the rigidly mapped spinal cord contour was 0.56 for this rCT). Furthermore, some part of the high-dose CTV moved slightly outside

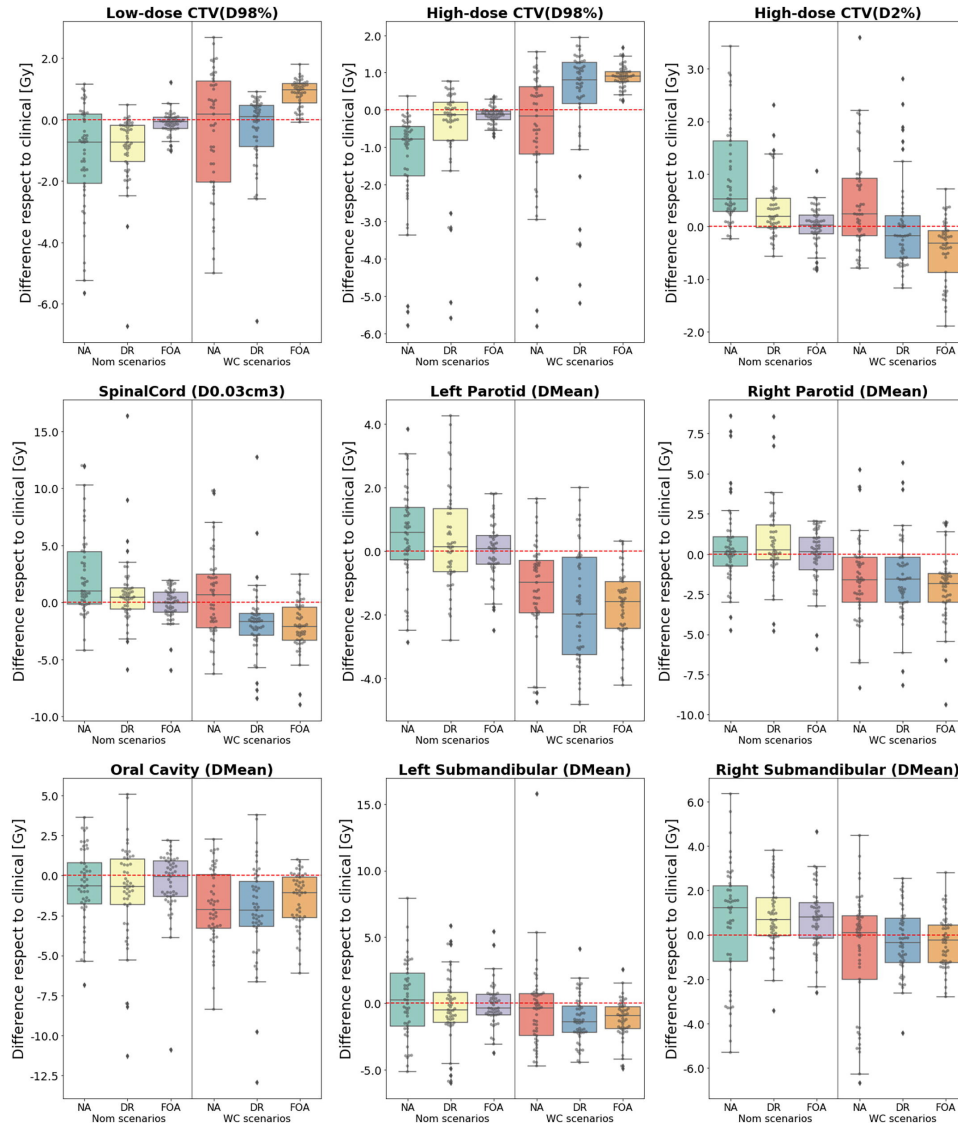


Figure 3.2: Differences in DVH-metrics with respect to clinical plan on the planningCT for not-adapted (NA), restored (DR) and fully offline adapted (FOA) plans for the nominal dose distribution (green, yellow and purple) and worst-case values (red, blue and orange) were reported for 50 rCTs. Abbreviations: $D98\%$ = dose received by 98% of the volume, $D2\%$ = dose received by 2% of the volume, $Dmean$ = mean dose, $D0.03\text{ cm}^3$ = dose received by a volume of 0.03 cm^3 , Nom = nominal, WC = worst-case.

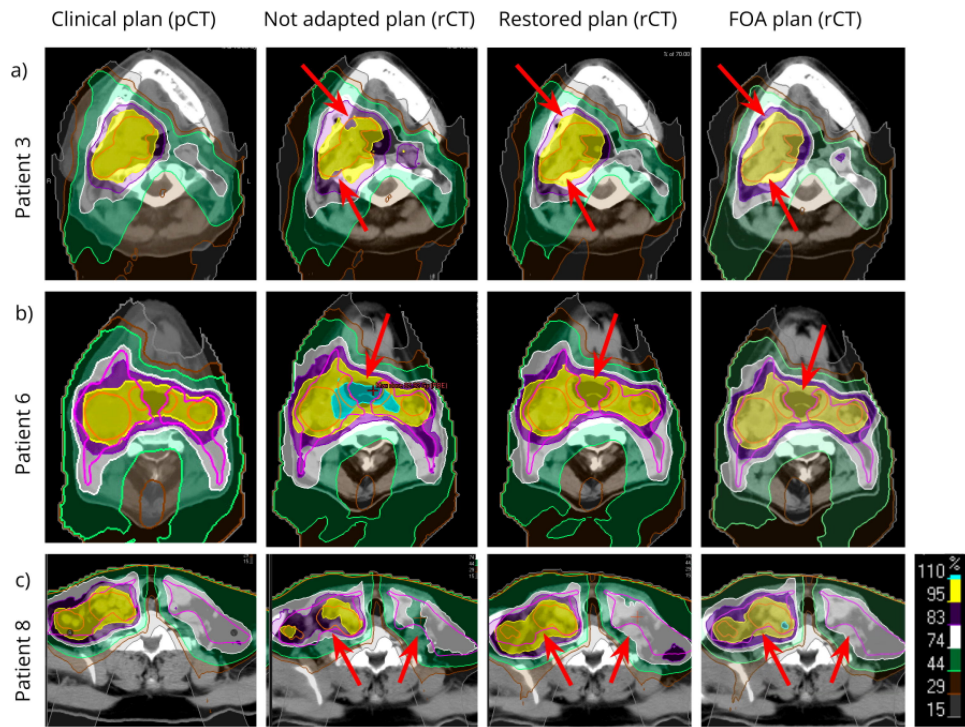


Figure 3.3: Nominal dose distributions in three representative patients for which dose restoration ensured target coverage but no adaptation did not. Red arrows highlight regions of over- or under-dosage. Isodoses are represented with different colors as a percentage of the prescribed dose, the yellow isodose is 95%(70 Gy) and the white isodose is 95%(54.25 Gy).

the high-dose region (the DSC score for the high-dose CTV was 0.74). For both these patients, a full offline adaptation was required to achieve an optimal target coverage.

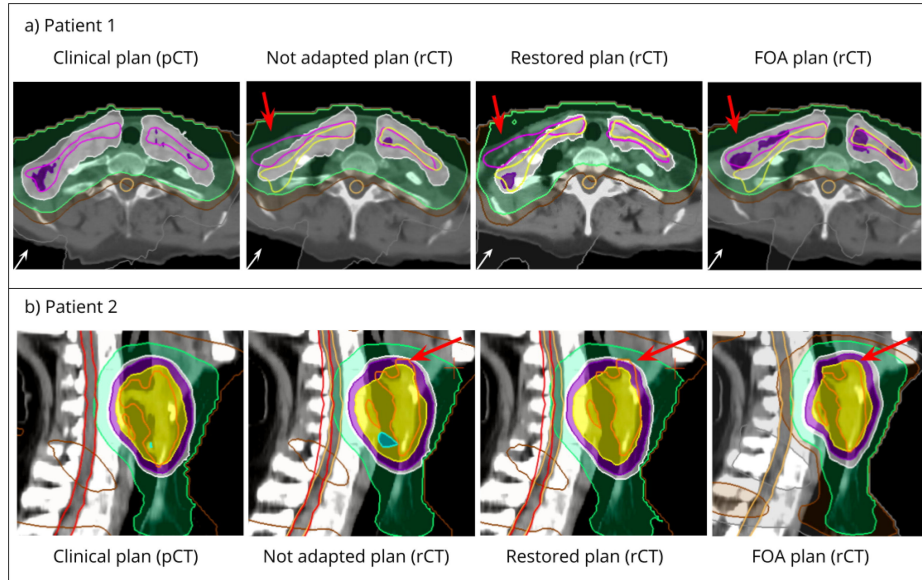


Figure 3.4: Nominal dose distributions in two representative patients where DR was not sufficient to ensure full target coverage. Red arrows highlight regions of over- or under-dosage. Isodoses are represented with different colors as a percentage of the prescribed dose, the yellow isodose is 95%(70Gy) and the white isodose is 95%(54.25Gy). a) The yellow contour is the rigidly mapped low-dose CTV and the magenta contour is the manually delineated low-dose CTV. b) The orange contour is the manually delineated high-dose CTV.

3.4 Discussion

Two adaptive strategies, (1) online dose restoration and (2) full offline adaptation, were compared in this study against not-adapted robustly optimized plans. Both strategies used robust optimization and aimed at mitigating dose distortions of IMPT plans induced by inter-fractional anatomical changes in a set of 50 rCTs for ten HNC patients.

It was shown that online dose restoration had the potential to reduce the need for offline adaptation of robustly optimized plans from 74% (37/50) to 48% (24/50). Three main findings can be extracted from our results.

First, robust optimization with the used parameters did not ensure adequate robustness levels against most anatomical changes observed in this study making the implementation of adaptive strategies indispensable. However, although many centers wish to implement adaptation workflows, they face barriers of hu-

man/material resources and technical limitations [160]. Dose restoration can be easily implemented for any tumor location and planning strategy. 4D-robust dose restoration was tested in lung cancer patients with a PROTEUS1 (IBA, Belgium) beam model [237] and thanks to the flexibility of the scripting module of RayStation, the workflow was adapted without difficulty to HNC patients treated with a MEVION beam model. Moreover, the simplicity of dose restoration allows its implementation in other treatment planning systems. Therefore, we believe dose restoration is accessible to any group interested in adaptive IMPT.

Second, as it has already been shown for other locations, such as lung tumors by Borderias et al. [237] or prostate by Jagt et al. [238], online dose restoration strategies using spot intensity re-optimization have the potential to automatically and successfully restore the clinical plan on rCTs. Hence, the clinical implementation of dose restoration could ensure target coverage for patients where robust treatment plans tend to be distorted. This result agrees with the results found by van de Water et al. [124], where superior target coverage and OAR sparing were shown for daily online dose restoration compared to robust optimization approaches in cases with an artificially generated nasal cavity filling.

Third, our study highlighted the intrinsic limitations of dose restoration which are especially remarkable in HNC. For a specific subset of patients, both robust and restored plans could not compensate for large anatomical changes which compromised the target coverage, making a cumbersome full offline re-planning still mandatory. Challenging shoulder repositioning seems to be especially critical for low-dose CTV coverage [240, 241] while the reduction of the neck diameter or drastic non-rigid deformations can affect both low and high-dose CTVs. Lalonde et al. [126] proposed an alternative online adaptation strategy using non-rigidly deformed contours which showed a higher efficiency with respect to classical and anatomical robust optimization. However, the use of this technique comes at the cost of the need for contour verification and doctor approval.

The selection of patients who would benefit from dose restoration is crucial since highly sensitive action levels towards offline adaptation may result in frequent adaptation with little clinical gain at the cost of high stress on human resources. Our study suggests potential DSC thresholds that could be defined for patient selection. It was found that low DSC coefficients (< 0.75) between rigidly mapped and manually delineated contours seemed to properly trigger cases where DR could not ensure sufficient target coverage. On the contrary, for high DSC coefficients (> 0.90) all restored plans had optimal target coverage (Appendix B, Figure B.2). However, further investigation is needed in a larger cohort of patients to assess the usability of DSC coefficient as a trigger for the use of DR. For intermediate

cases ($0.75 < \text{DSC} < 0.90$), a fast-independent dose calculation could also help to select the appropriate adaptation strategy and as a pre-treatment patient-specific QA procedure. Since the manually delineated contours are not available immediately after a rCT is acquired, deformable image registration or automatic segmentation tools [242] could be used in this evaluation step. Other parameters such as the degree of weight loss and reduction in lateral neck diameter could also be monitored as predictors of patients receiving the greatest benefit of adaptation [243]. In any case, online adapted plans should be followed by an extensive and fast, patient-specific quality assurance including machine log files after delivery [104] together with independent Monte Carlo dose engine calculations [154, 244, 245].

As previously mentioned, the two key points of dose restoration are (1) no need for manual re-contouring and (2) maintenance of unchanged dose distribution on rCTs to minimize the physician's approval step. However, we have seen that there are limitations to this approach. Emerging commercial solutions for conventional radiation therapy using CBCT based workflows can already create plans with completely new dose distributions adapted to a new anatomy in a few minutes [194]. Similarly, an ideal IMPT adaptive workflow could use, for example, deep learning models for automatic contour segmentation [246, 247] followed by dose prediction [248, 223] and a dose mimicking optimization algorithm [35, 249]. However, quality assurance processes required to verify completely new plan designs may make online adaptation cumbersome for clinical implementation.

A re-computation of the final doses with a smaller dose grid ($2.0 \times 2.0 \times 2.0 \text{ mm}^3$) is recommended to not underestimate the values of DVH metrics in smaller OAR, such as the optical structures. The dose to these organs can be a limiting factor to achieve the prescribed dose to the target, especially for nasopharynx tumors. Although this was not an issue in our study, it could be for many other nasopharyngeal cases.

In our study, the total time for plan adaptation ranged between 11.9 and 23.1 min, with a median of 17.4 min. The median time distribution was 6.2 min for the first step including pCT-rCT rigid registration, contour mapping and loading spot information and 11.2 min for spot intensity robust re-optimization. This time could be different for other treatment machines than the proton therapy and dose calculation algorithm used in our study (MEVION S250i) which requires Monte Carlo dose computation of all error scenarios during robust optimization due to the use of the multi-leaf collimator. Using interpolation during optimization in RayStation can reduce the optimization time by a factor of three. In any case, the time lost was purely computational, and therefore its reduction should be

addressed via hardware investment or the improvement of optimization algorithms.

In conclusion, automatic online dose restoration could reduce the offline adaptation rate in HNC IMPT treatments, saving valuable human and material resources to radiotherapy departments. Unfortunately, full offline adaptation was still necessary for 48% of the cases and only a reduced group of patients suffering specific density changes could benefit from dose restoration.

Chapter 4

Robustness in automatic IMPT planning: impact of post-processing and dose-mimicking

The content presented in this chapter comes from a manuscript submitted to Medical Physics under the title: "*Machine learning-based automatic proton therapy planning: impact of post-processing and dose-mimicking in plan robustness*" and which is at the moment under revision. This work was carried out in close collaboration with the Machine Learning department from RaySearch Laboratories, Stockholm (Sweden).

Authors: Elena Borderias-Villarroel¹, Margerie Huet¹, Ana Barragán-Montero¹, Rasmus Helander², Mats Holmstrom², Xavier Geets^(1,3) and Edmond Sterpin^(1,4)

¹ UCLouvain, Institut de recherche expérimentale et clinique, Molecular Imaging and Radiation Oncology Laboratory, Brussels, Belgium. ² RaySearch Laboratories, Stockholm, Sweden. ³ Cliniques Universitaires Saint Luc, Brussels, Belgium. ⁴ KULeuven, Department of Oncology, Laboratory of external radiotherapy, Leuven, Belgium.

Summary

Background and purpose: Automated treatment planning strategies are being widely implemented in clinical routine to reduce inter-planner variability, speed up the optimization process, and improve plan quality. This study aims to evaluate the feasibility and quality of intensity-modulated proton therapy (IMPT) plans generated with four different knowledge-based planning (KBP) pipelines fully integrated into a commercial treatment planning system (TPS).

Methodology: A data set of 60 oropharyngeal cancer patients was split into 11 folds, each containing 47 patients for training, 5 patients for validation and 5 patients for testing. A dose prediction model was trained on each of the folds, resulting in a total of 11 models. Three patients were left out in order to assess if the differences introduced between models were significant. From voxel-based dose predictions, we analyzed the two steps that follow the dose prediction: post-processing of the predicted dose and dose mimicking (DM). We focused on the effect of post-processing (PP) or no post-processing (NPP) combined with two different DM algorithms for optimization: the one available in the commercial TPS RayStation (RSM) and a simpler isodose-based mimicking (IBM). Using 55 test patients (5 test patients for each model), we evaluated the quality and robustness of the plans generated by the four proposed KBP pipelines (PP-RSM, PP-IBM, NPP-RSM, NPP-IBM). After robust evaluation, dose-volume histogram (DVH) metrics in nominal and worst-case scenarios were compared to those of the manually generated plans.

Results: Nominal doses from the four KBP pipelines showed promising results achieving comparable target coverage and improved dose to organs at risk (OARs) compared to the manual plans. However, too optimistic post-processing applied to the dose prediction (i.e. important decrease of the dose to the organs) compromised the robustness of the plans. Even though RSM seemed to partially compensate for the lack of robustness in the PP plans, still 65% of the patients did not achieve the expected robustness levels. NPP-RSM plans achieved the best trade-off between robustness and OAR sparing.

Conclusions: PP and DM are crucial steps to generate acceptable robust and deliverable IMPT plans from ML-predicted doses. Before the clinical implementation of any KBP pipeline, the PP and DM parameters predefined by the commercial TPS need to be modified accordingly with a comprehensive feedback loop in which the robustness of the final dose calculations is evaluated. With the right choice of PP and DM parameters, clinically acceptable IMPT plans could be generated.

4.1 Introduction

Radiation therapy (RT) treatment planning in clinics is currently a time-consuming process involving tedious manual operations. Automated treatment planning workflows are being widely implemented in clinical settings to reduce inter-planner variability, speed up the optimization process, and improve plan quality [250].

Various approaches have recently been developed that use prior knowledge (e.g., knowledge from libraries of previously treated patients) to automate treatment planning (knowledge-based planning, or KBP)[251]. In one of its forms, KBP can be described as a two-stage pipeline. In the first stage, a trained machine learning (ML) model uses binary volumes of the regions of interest (ROIs) to predict the dose that should be delivered to each voxel of the patient. In the second stage, the dose prediction from the first stage is used in an optimization process to generate a deliverable treatment plan. After the user has chosen a beam configuration, the optimization process begins. The optimization step is crucial because it determines the machine parameters required to deliver the desired dose distribution. In the case of intensity-modulated proton therapy (IMPT) delivered with pencil beam scanning (PBS), the optimization will compute the number of energy layers needed to cover the tumor in-depth, the spot positions for each layer, and the spot weights for each beam.

Dose mimicking (DM) is one of the optimization processes that can follow the dose prediction step in the KBP pipeline [35, 249]. The optimization used for DM is the same as in regular treatment planning, with algorithms like gradient descent for optimization and analytical or Monte Carlo methods for dose calculation. Babier et al. [252] compared different combinations of two prediction models (a generative adversarial network and a random forest) and two optimization models (inverse planning and dose mimicking) in head and neck IMRT. According to the findings of this research, the performance of an automated KBP pipeline is determined by how well the prediction and optimization models work together.

Significant research has been conducted to advance KBP, with a primary emphasis on improving ML models like ResNet or UNet [253, 222, 254]. Most of the research has focused on photon radiation therapy techniques such as intensity-modulated radiation therapy (IMRT) and volumetric modulated arc therapy (VMAT), with only a few studies addressing proton therapy in the pelvic [255, 256] and abdominal [248] regions.

Optimization models, on the other hand, have received far less attention in the literature than dose prediction models. They are, however, particularly important in proton therapy to ensure the robustness of the generated treatment plans. Proton dose distributions are more sensitive to treatment uncertainties than photon doses [15, 16]. This means that the robustness of proton dose distributions is not solely dependent on the dose shape, rendering the margin strategy suboptimal for IMPT. Robust optimization strategies have been proposed as a more appropriate strategy for the generation of high-quality proton plans [77, 72, 73, 257, 258].

Automated planning workflows are not always flawless. An ML model may predict suboptimal dose distributions, such as an overdose of an organs at risk (OAR), insufficient target coverage of the clinical target volume (CTV), or an unrealistic trade-off between OAR sparing and CTV coverage. These effects could be compensated for by post-processing the predicted dose and/or tweaking the setup of the dose mimicking problem to improve the deliverability of the final dose. The commercial treatment planning system (TPS) RayStation 11B includes an ML module for automatic IMPT treatment planning including both post-processing and dose mimicking. Before the plan optimization, an in-built post-processing step with pre-configured clinic-specific settings for each protocol runs in their KBP pipeline. The post-processing is applied to the predicted dose aiming to ensure enough target coverage and improve OAR sparing. Such post-processing, together with the configuration of the dose-mimicking parameters, are crucial to the KBP pipeline. This is particularly important for KBP pipelines in proton therapy, since post-processing and mimicking parameters might play an important role in the robustness of the generated IMPT plan. The robustness of a similar KBP pipeline introducing several predicted-scenario doses was explored by Eriksson and Zhang for prostate patients [258]. However, to our knowledge, no study has evaluated the combination of a post-processing step with dose mimicking so far, while it seems vital before ML-based KBP pipelines can be implemented in the clinic. This is especially important in head and neck regions, where the presence of multiple nearby OAR that are preferably spared as much as possible makes the implementation of these automated planning strategies even more complex and challenging.

In this paper, we studied four complete KBP pipelines for IMPT plan generation in oropharyngeal cancer. The steps following dose prediction were the focus of this research. We investigated the effect of post-processing (PP) or no post-processing (NPP) on the final dose distribution when combined with two different dose mimicking algorithms: the one available in RayStation (RSM) and a simpler isodose-based mimicking (IBM). Overall, the four proposed KBP pipelines (PP-RSM, PP-IBM, NPP-RSM, and NPP-IBM) were built and benchmarked against

the corresponding clinical plans using DVH-based metrics. The analysis of the robustness of the final treatment plans in the four settings is also given special attention.

4.2 Material and methods

4.2.1 Patient cohort and planning strategy

Our dataset included 60 patients with bilateral oropharyngeal cancer which were manually planned by the same dosimetrist using a treatment planning protocol to ensure consistency (Appendix C, SM1, Table C.1). Local ethics committee approval was obtained for the use of all patient data. The prescribed dose (D_p) was 70 GyRBE to the high-dose CTV and 54.25 GyRBE to the low-dose CTV in 35 fractions with a constant relative biological effectiveness (RBE) factor of 1.1. The parotids, submandibular glands (SMGs), the pharyngeal constrictor muscles (PCMs), the oral cavity, the spinal cord and the brainstem were defined as the main OAR. Target and OAR delineations performed by experienced radiotherapy oncologists were available on all planning-CTs.

The same beam configuration was used for all patients: two anterior beams at $(60^\circ, 10^\circ)$, and $(120^\circ, 10^\circ)$ and two posterior beams at $(240^\circ, 350^\circ)$ and $(300^\circ, 350^\circ)$ for gantry and couch angles, respectively. For each patient, a robustly optimized treatment plan was created using Monte Carlo dose calculation and classical worst-case robust optimization (setup error (SE) = 4 mm, range error (RE) = 2.6%). Although 3% of range uncertainty is generally used in the clinic, when applying the 1.6% sigma to range uncertainty, as quantified by Paganetti [18], to Van Herk's margins formula at a 90% confidence level [66], the resulting range uncertainty can be estimated as $1.6\% * 1.64 = 2.6\%$.

SE scenarios included patient shifts along the three space axes x, y, z ($\pm SE, 0, 0$), $(0, \pm SE, 0)$, $(0, 0, \pm SE)$, and the nominal scenario $(0, 0, 0)$. Range robustness was also included to account for uncertainties in the conversion of CT numbers to proton stopping power ratios. RE scenarios were modelled by scaling the mass density by $\pm 2.6\%$. The nominal scenario (i.e. 0%) was also included. The combination of seven SE scenarios and three RE scenarios resulted in a 21 error scenario optimization (7 SE scenarios $\times$ 3 RE scenarios). Robust optimization was followed by a robust evaluation using the same robustness parameters (SE = 4 mm, RE = 2.6%). The robust evaluation was performed by simulating 42 uncertainty scenarios (14 SE scenarios $\times$ 3 RE scenarios). According to Korevaar et al. [101], the 6 principal directions ($\pm SE, 0, 0$), $(0, \pm SE, 0)$, $(0, 0, \pm SE)$, $(0, SE, 0)$ and the 8 diagonals from the center to the vertices ($SE/\sqrt{3}, SE/\sqrt{3},$

$SE/\sqrt{3}$) in combination with 0, RE% were the sampled error scenarios. Plans were manually adjusted until they met the list of clinical goals as closely as possible in both the nominal and worst-case scenarios. The resulting plan meeting the clinical goals included in the planning protocol was considered the ground truth of the study and will be referred to as the clinical plan from this point.

4.2.2 Dose prediction model

Spatial dose prediction is available in RayStation 11B using a UNet architecture that takes as input 3D binary matrices from the delineated volumes. The network is trained in a supervised manner to learn how to predict a dose value for each voxel. More information concerning the specifications of the prediction model can be found in Çiçek et al. [259]. In order to get the most out of the patient dataset containing 60 patients, 11 dose prediction models were trained using 47 patients for training, 5 patients for validation, and 5 patients for testing. Three patients (check set) were left out to determine whether the differences introduced by the different models were significant or negligible. The partition of the dataset is explained in Figure 4.1.

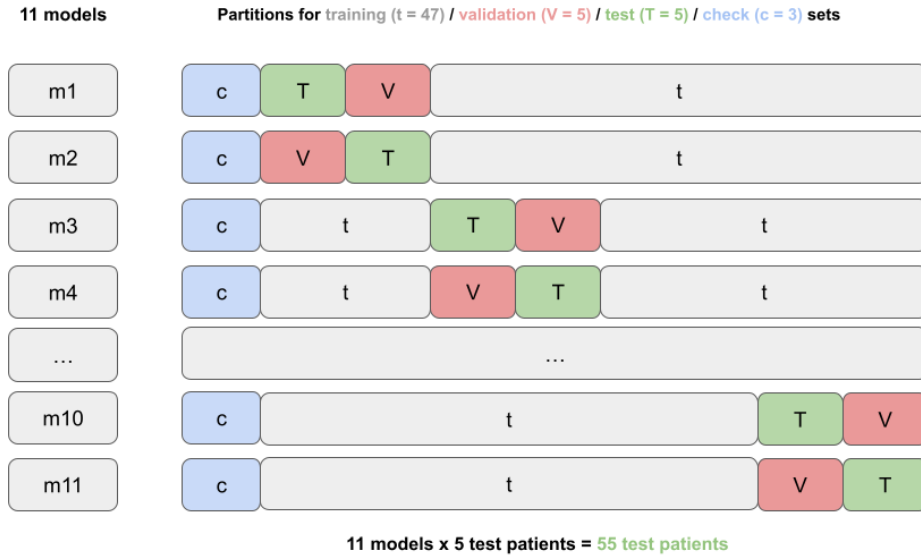


Figure 4.1: Database partitioned into training, validation, test, and check sets. The check set consists of 3 left-out patients to assess the difference between the 11 generated models. With this dataset partition a total of 55 test patients were obtained for this study. Abbreviations: m (model), c (check set), T (test set), V (validation set), t (training set).

4.2.3 Post-processing of the prediction

Before setting up the dose mimicking problem, the predicted spatial dose distribution of the UNet model goes through a post-processing (PP) phase. This phase is designed to ensure that the deliverable (mimicked) dose adheres to a set of pre-defined clinical goals. The post-processing is implemented in RayStation as a set of ROI-specific alterations of the DVHs of the predicted dose distribution, after which the spatial dose is resampled to fit the new DVHs. This step serves to drive the dose mimicking optimization towards a deliverable dose which is more clinically preferable than the raw UNet prediction.

The parameters of the post-processing step were chosen to be the same as the ones from the ML model 'RSL-IMPT-Oropharynx-7000-SIB (3.0)', which was trained and configured for a clinical protocol similar to the one of this study [260]. The parameters are presented in Appendix C (SM2, Table C.2) and their detailed specifications can be found in the guide "RayStation 11B – Scripting Environments for Machine Learning". Once the post-processing settings are chosen for a specific clinic and protocol, the ML model will then be available for use in all subsequent patients without the need for manual tuning.

4.2.4 Dose mimicking strategies

The information of the predicted dose (with or without PP) was given as the dose to mimic to two different dose mimicking strategies. The IBM uses isodose volumes generated from the dose to mimic while RSM aimed at minimizing the difference with the dose to mimic by including voxel-wise-based functions as optimization objectives. The two mimicking methods used the RayStation worst-case robust optimization (see the Appendix C, SM3). The robust functions used in RSM and IBM are conventional objective functions such as minimum dose and maximum dose to the target volumes. The robustness parameters were 2.6% for the range error and 4 mm for the setup error as described in section 4.2.1. Both methods were fully integrated into the commercial TPS RayStation 11B. The implementation and parameters of each mimicking method are described in the next paragraphs:

Isodose based dose mimicking

Isodose Based Mimicking (IBM) used isodose volumes generated from the predicted dose every 2 Gy within 95%–107% of the prescription dose and 5 Gy outside as suggested by Bernatowicz et al. [33] for isodose volume dose restoration. Optimization was performed using minimum and maximum dose objectives on the CTVs and on each isodose volume.

Maximum and minimum dose objectives were set for isodoses above 70% of the prescribed dose, while isodoses below 70% of the prescribed dose only had a maximum objective. This allows placing the spots in a restricted area around the CTV to prevent the optimizer from setting spots in the whole patient volume. A total of 11 dose objectives were set as robust in order to ensure target coverage and avoid hot spots caused by treatment uncertainties. Contrary to other studies using IBM where the optimization objectives are patient-specific [33, 261], a single optimization template was generated. The weights of the objectives were empirically chosen after manual tuning in the subset of the three left-out patients taking as a starting point the experience described in [261]. The final parameters were saved as the template that was used later to run the optimization (i.e. dose mimicking) for all patients. The IBM optimization objectives are detailed in the Appendix C (SM4, Table C.3). IBM took around 10 minutes to complete the optimization step for the 21 robust scenarios.

RayStation dose mimicking

RayStation Mimicking (RSM) is the dose mimicking method integrated in the ML model 'RSL-IMPT-Oropharynx-7000-SIB (3.0)'. It is formulated as a voxel-wise least squared error (LSE) minimization problem. The voxel weights are determined based on the geometric position of the voxel in relation to the ROIs, as well as by the assigned dose value of the voxel in the post-processed dose distribution. The LSE objective function is divided into two one-sided terms, enabling different weights depending on whether the optimized dose is greater or less than the post-processed reference dose – like the max and min dose strategy in IBM. For OAR, voxels with dose values lower than the reference dose are not penalized. In addition to the voxel-wise LSE objective, an additional term is added to the optimization problem. This term consists of a weighted sum of conventional radiotherapy objective functions, as implemented in RayStation 11B.

Similar to the post-processing parameters, the settings of the dose mimicking optimization only need to be configured once per ML Model. For this study, these settings were based on the 'RSL-IMPT-Oropharynx-7000-SIB (3.0)' model, with a change in the relative importance of the brainstem. Details on the used parameters are found in Appendix C (SM5, Table C.4). The optimization procedure consists of two separate runs, between which the clinical dose is calculated accurately. These two runs consist of 125 and 80 iterations respectively, and the full optimization procedure is completed in approximately 20 minutes for the 21 robust scenarios.

4.2.5 Automatic planning strategies

In collaboration with RaySearch, 11 prediction models were generated and imported into the ML module. The four KBP pipelines (PP-RSM, PP-IBM, NPP-RSM, and NPP-IBM) shown in Figure 4.2 were fully integrated into RayStation. They could be easily run by the user with a single mouse click. PP-RSM and NPP-RSM could also be launched directly from RayStation's GUI, while scripting was necessary for PP-IBM and NPP-IBM methods. The same robustness evaluation described in 4.2.1 for the clinical dose was applied to the dose distributions automatically generated using the proposed KBP pipelines. All dose calculations were run in RayStation 11B, using the Monte Carlo algorithm on a $3.0 \times 3.0 \times 3.0$ mm³ dose grid.

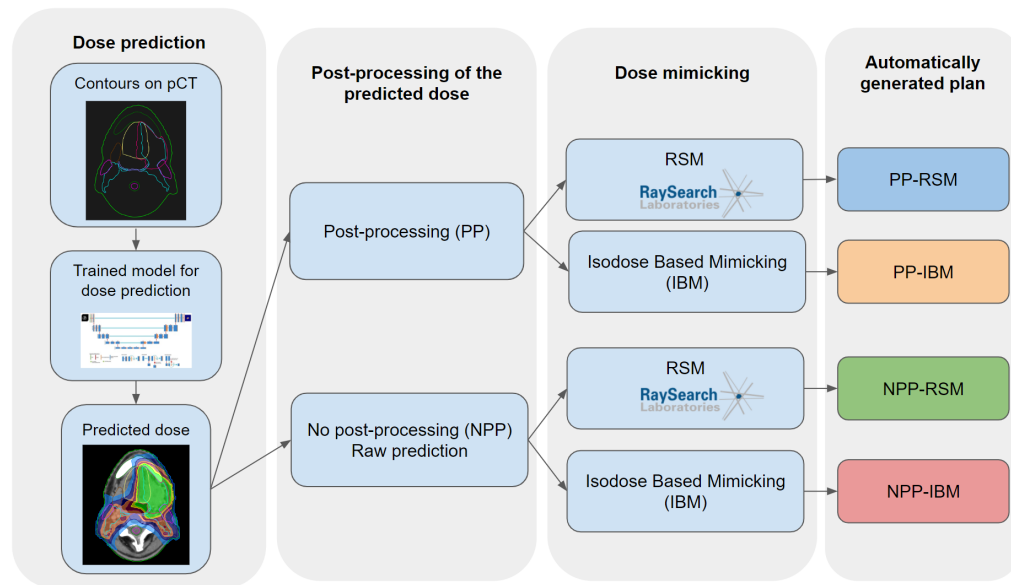


Figure 4.2: Overview of the automated knowledge-based planning pipelines evaluated in our study.

4.2.6 Evaluation metrics

Clinically relevant DVH-parameters of OAR and CTVs were compared between the four KBP pipelines and the clinical plan. Differences with respect to the clinical dose in D98 values (dose received by 98% of the volume) were calculated for low-dose and high-dose CTVs, while differences in Dmean (average dose) were reported for the OAR. These metrics, evaluated on the 55 patients, were displayed

using boxplots. The median and interquartile range (IQR) values from the boxplots illustrate how much these metrics differ from the clinical dose for each KBP method. For low-dose and high-dose CTV, nominal scenario (with no uncertainty) and worst-case scenario values were calculated from the 42 uncertainty explored scenarios ($SE = 4\text{mm}$, $RE = 2.6\%$) during robustness evaluation as described in section 2.1. Only nominal values for the OAR average dose were reported.

Grouped data were tested for statistical significance using two-sided Wilcoxon-Mann-Whitney tests using SciPy version 1.5.2. The differences were considered statistically significant for p-values < 0.05 .

4.3 Results

4.3.1 Nominal target volume coverage

For the four KBP pipelines, the clinical goals for target coverage (i.e., $D98\% \geq 95\%$) in both high-dose and low-dose CTVs were met for all the patients in the nominal scenario. Figure 4.3 shows the difference in DVH-metrics between the clinical dose and the four investigated methods. High-dose CTV target coverage decreased by less than 0.50 Gy in median for the four strategies. PP-IBM plans presented the largest decrease (i.e. lowest D98 values) with a reduction in the median value of -0.42 (iqr = 0.27) Gy. NPP-RSM plans achieved the best coverage with an increase in the median value of 0.05 (iqr = 0.20) Gy with respect to the clinical dose. For the low-dose CTV, D98 values were closer to no difference or even higher than clinical D98 values. All KBP methods had less than a 0.12 Gy reduction in the median target coverage. The difference in D98 values for the low-dose CTV between the PP-IBM and clinical plans was not statistically significant.

4.3.2 Worst-case target volume coverage

In contrast to the nominal scenario, in the worst-case scenario, not all the clinical goals were met for all patients. Especially in the high-dose CTV, plans where post-processing was applied, PP-RSM and PP-IBM, experienced a significant drop in target coverage with -1.75 (iqr = 1.24) Gy and -2.35 (iqr = 1.65) Gy, respectively. 65% (35/55) and 82% (45/55) of patients failed to achieve high-dose CTV coverage in the worst-case scenario for PP-RSM and PP-IBM, respectively. Without post-processing, only 5% (3/55) and 11% (6/55) of patients failed to meet the robustness criteria for NPP-RSM and NPP-IBM, respectively. A patient example of the robustness loss in the high-dose CTV can be found in Appendix C (Figure C.1). While all scenarios achieved sufficient high-dose CTV coverage

in the initial clinical plan, NPP-RSM and NPP-IBM plans, only 83% and 31% of the 42 evaluated scenarios provided enough target coverage for PP-RSM and PP-IBM, respectively.

A higher robustness level was observed for the low-dose CTV. All patients achieved target coverage in the worst-case scenario for PP-RSM and NPP-RSM plans. For PP-RSM, there was no statistically significant difference in the worst-case scenario values with respect to clinical plans. However, PP-IBM and NPP-IBM did not meet adequate low-dose CTV coverage in the worst-case scenario in 5% (3/55) and 10% (6/55) of the patients.

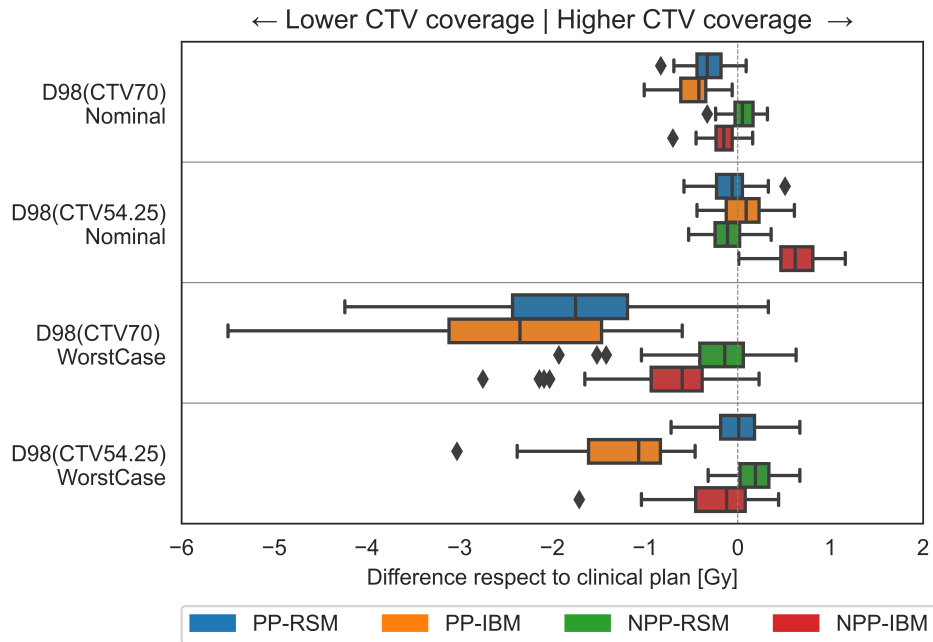


Figure 4.3: The differences in target coverage (D98 - dose received by 98% of the CTV) between the evaluated KBP pipelines and the clinical plans in nominal and worst-case scenarios. Negative differences imply that the clinical plan was better than the KBP method for that metric. The boxes indicate the median and interquartile range (IQR). The whiskers extend to a minimum of 1.5 times the IQR, and diamonds represent the outliers. Abbreviations: KBP = Knowledge-Based Planning, PP = Post-processing, NPP = No Post-Processing, IBM = Isodose-Based Mimicking, RSM = RayStation Mimicking.

Clinical goal D98(CTV) <95%Dp	Scenario	Failure of the clinical goal for each plan: absolute number (% of total 55 patients)			
		PP-RSM	PP-IBM	NPP-RSM	NPP-IBM
CTV70	Nominal	0	0	0	0
CTV54.25		0	0	0	0
CTV70	Worst-case	35 (63%)	45 (82%)	3 (5%)	6 (11%)
CTV54.25		0	16 (29%)	0	1 (2%)

Table 4.1: For each of the evaluated KBP pipelines, nominal and worst-case scenarios, the number of patients out of the 55 that fail at achieving clinical goals. The number of patients is reported in absolute value and as a percentage of the total 55 patients. The KBP pipelines using post-processing compromise the robustness of the high-risk CTV in more than 50% of the cases.

4.3.3 Nominal dose to organs at risk

Figure 4.4 presents the difference in the average dose received by the OAR between the clinical dose and the four studied KBP methods. For all OAR metrics, the differences with respect to the clinical plan and between methods were statistically significant. In most of the cases, the plans generated with KBP pipelines achieved superior OAR sparing than the clinical plans. Median values of the average doses were lower than in the clinical plan for most OAR. The few exceptions were the esophagus in NPP-RSM plans and the SMGs in NPP-IBM plans. As expected, OAR sparing on the oral cavity, the parotids, and PCMs was especially aggressive for plans including post-processing according to the pre-defined settings. PP-RSM and PP-IBM plans saved with respect to the clinical plan median values of more than 6 Gy to the oral cavity, more than 5 Gy to the parotids, and between 4.5 and 5 Gy to the PCMs. An illustrative example is shown in Figure 4.5.

The number of patients failing at achieving OAR clinical goals with a higher value than the clinical plan is reported in Table 4.2. Some clinical goals were very difficult to be respected even for the initial clinical plan. When comparing the results of the KBP pipelines, the clinical goals were met in most of the cases, or at least the dose to OAR was improved with respect to the clinical plan, especially in PP-RSM and PP-IBM plans. In the same way, as observed in Figure 4, NPP-IBM was the KBP pipeline with higher doses to OAR being the SMGs and the superior and medium PCMs the most challenging to be spared.

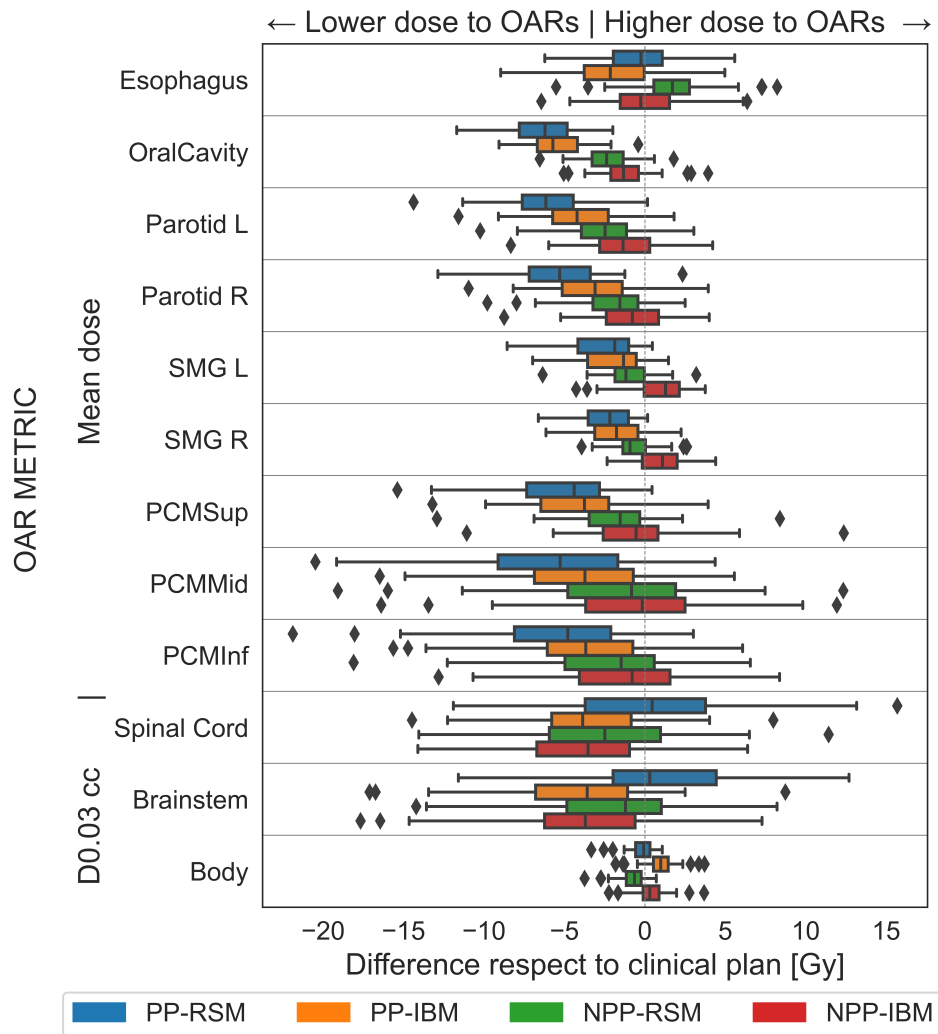


Figure 4.4: The differences in terms of clinical planning criteria for the OARs between KBP plans and the clinical plans. Negative differences imply that the KBP plan was better than the clinical plan for that metric (i.e., less dose to the OARs). The boxes indicate the median and interquartile range (IQR). The whiskers extend to a minimum of 1.5 times the IQR and diamonds represent the outliers. Abbreviations: KBP = Knowledge-Based Planning, PP = Post-processing, NPP = No Post-Processing, IBM = Isodose-Based Mimicking, RSM = RayStation Mimicking, SMG = Submandibular Gland, PCM = Pharyngeal Constrictor Muscle

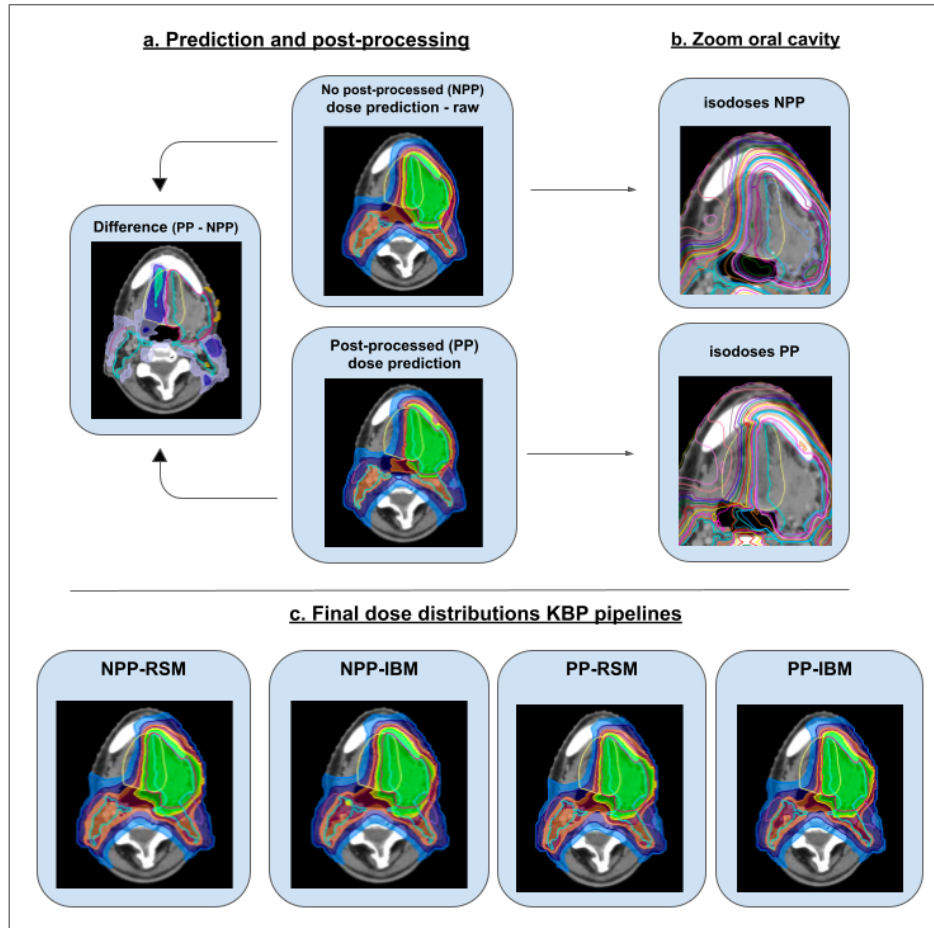


Figure 4.5: Different steps of the proposed KBP pipelines in a patient example. The image on the left (a) shows the differences between the raw predicted dose (NPP) and the post-processed predicted dose (PP): light blue (- 4 Gy), dark blue (-7 Gy), and turquoise blue (-15 Gy). In the middle (b) the isodose contours generated from the PP and NPP doses that are used during isodose based mimicking are presented. Specifically, a zoomed image of the oral cavity region is shown. On the left (c), the final result of the four proposed KBP workflows for this patient can be seen. The delineated contours are high-dose CTV (magenta), low-dose CTV (turquoise blue) and the oral cavity (yellow). The isodoses characterizing the dose distributions are represented with different colours as a percentage of the prescribed dose of 70 Gy: light blue (25%), dark blue (50%), orange (74%), dark red (81%), yellow (90%), green (95%) and red (105%). The green isodose corresponds to 95%(70 Gy) and the white isodose is 95%(54.25 Gy).

Clinical goals for the organs at risk	Priority	Number of patients out of 55 that failed at achieving the clinical goal (% of total 55 patients)				
		Initial plan	KBP pipelines			
			PP-RSM	PP-IBM	NPP-RSM	NPP-IBM
D0.03cc(BrainStem) >54 Gy	1	0	0	0	0	0
D0.03cc(SpinalCord) >45 Gy	1	0	0	0	0	0
Dmean(PCMSup) >40 Gy	3	54 (98.2%)	0	3 (5.4%)	6 (10.9%)	10 (34.5%)
Dmean(PCMMid) >40 Gy	3	31 (56.4%)	1 (1.8%)	2 (3.6%)	13 (23.6%)	16 (29.1%)
Dmean(PCMInf) >40 Gy	3	10 (18.2%)	0	0	1 (1.8%)	0
Dmean(Parotid L) >26 Gy	3	27 (49.1%)	0	0	1 (1.8%)	5 (9.1%)
Dmean(Parotid R) >26 Gy	3	20 (36.4%)	0	1 (1.8%)	3 (5.4%)	7 (12.7%)
Dmean(SMG L) >40 Gy	3	51 (92.3%)	4 (7.3%)	10 (18.2%)	10 (18.2%)	37 (67.3%)
Dmean(SMG R) >40 Gy	3	47 (85.5%)	2 (3.6%)	7 (12.7%)	9 (16.4%)	34 (61.8%)
Dmean(Oral Cavity) >35 Gy	3	36 (65.5%)	0	0	2 (3.6%)	6 (10.9%)
Dmean(Esophagus) >35 Gy	3	3 (5.5%)	0	0	5 (9.1%)	2 (3.6%)

Table 4.2: For each of the evaluated KBP pipelines, the number of patients out of the 55 that fail at achieving the organ at risk clinical goals with a higher value than the initial clinical plan. Reported values correspond to the evaluation of clinical goals only on nominal dose distributions. Abbreviations: KBP = Knowledge-Based Planning, PCM = Pharyngeal Constrictor Muscles, SMG = Submandibular Glands

4.4 Discussion

Automated KBP pipelines for IMPT planning necessitate a comprehensive evaluation of all the steps included in the workflow for a safe clinical implementation [252]. In this study, the clinical IMPT plans of 55 oropharyngeal patients were compared to automated IMPT plans generated using four different KBP pipelines (PP-RSM, PP-IBM, NPP-RSM, and NPP-IBM). The impact of a post-processing strategy combined with two different robust dose mimicking strategies was analyzed. These two steps of the KBP pipelines, coming after the dose prediction (post-processing and dose mimicking), have not been fully addressed in the literature, yet they are essential to judge robustness and the clinical acceptability of automatically generated IMPT plans.

The four proposed KBP pipelines, generated nominal dose distributions meeting all the clinical goals. In many cases, they even achieved superior quality in the OAR sparing compared to manually planned clinical doses, especially both KBP pipelines that use post-processing before dose mimicking: PP-RSM and PP-IBM. The potential of reduction of the dose to OAR in HNC patients is especially important in order to reduce radiation-induced side effects, in particular xerostomia and dysphagia, which have a major impact on quality of life [13, 49].

On the other hand, a dosimetric gain in the OAR can come at the cost of losing robustness. The results show that in some cases, our post-processing parameters compromise the robustness of the plans, especially for the high-dose CTV. This is

probably due to the reduction of the dose to the OAR during the post-processing step. In particular, the dose reduction of 20% to the oral cavity from the predicted value, combined with the push for target homogeneity in the high-dose CTV, generates unrealistic dose gradients where both regions of interest meet (Figure 5a. and 5c.). These unrealistic dose gradients are then propagated to the isodose volumes, which end up being very close to each other (Figure 5b.). In such cases, the optimization problem becomes extremely challenging for PP-IBM plans. IBM could not compensate for the loss in robustness in the high-dose CTV but was able to maintain correct target coverage for the low-dose CTV, where the dose gradients were not so steep. A more sophisticated mimicking process, such as RSM, seemed to partially recover the robustness. However, PP-RSM still failed to satisfy high-dose CTV robustness in 65% of the cases. The lack of robustness in the target coverage could be translated into cold spots (i.e. underdosage) during the treatment delivery, which could lead to a drastic reduction of the tumor control probability [262]. In an automatic planning study by Kierkels et al. dose mimicking is combined with a reduce-dose step to generate a robust proton plan from a previously optimized photon dose distribution [263]. Robustness in their automatic planning pipeline failed when the dose to mimic presented an underdosed target volume. In our KBP, it is not the presence of an underdosage in the dose to mimic (predicted dose) but rather an unrealistic dose gradient yielding a not feasible physical dose due to the post-processing step.

It was shown that the construction of a KBP pipeline, including post-processing and dose mimicking, is an iterative process that requires adjustment of the settings provided by the commercial solution. With the current parameters of our study, NPP-RSM plans seem to achieve the best trade-off between robustness and OAR sparing. They present comparable target coverage as in the reference clinical plan and higher OAR sparing, except for the esophagus. In order to improve the generated plans using NPP-RSM, specific objectives aiming at reducing the dose to the esophagus could be introduced in the mimicking step. Another interesting option would be to adjust the post-processing parameters for PP-RSM plans to maximize the potential of the OAR sparing.

It should be pointed out that the post-processing parameters were not configured specifically for this study, but were taken from another (similar) model which might have been less strict in the desired robustness. It is likely that one would find a good set of post-processing parameters for this study as well, if iterated a few times. Therefore, it would be worth exploring if less aggressive post-processing can maybe still improve OAR sparing while not compromising so much on the robustness of the plans. In any case, these results highlight the difficulty of

generating universal KBP strategies (combining dose prediction, post-processing and dose mimicking) that are interchangeable between clinical protocols, even if they are similar. Although the same prediction model could be used, the post-processing and dose mimicking parameters would highly rely on specific clinical preferences.

The only input to the ML dose prediction used in our proposed KBP pipelines was the contours. Recent studies about dose prediction for radiotherapy have explored the use of auxiliary information (e.g., non-modulated beam doses) to improve the robustness of the ML model [223, 264]. Assuming that the robustness of a dose distribution relies on the IMPT plan, another option would be to directly predict those plan parameters (e.g. energy layers, spot positions and spot weights). Approaches suggesting the latter have been described in the literature using fluence map prediction to generate photon plans [265, 266, 267] but still need to be investigated for protons. However, even if ML models are excellent methods for fast, automatic, and free-of-any human intervention tasks, they are also well known for their low interpretability [185]. For the moment, there is no evidence suggesting that ML methods could also outperform the already available, fast, automatic and well-studied analytical computing models for specific tasks like dose calculation or treatment plan optimization. Therefore, with the current technology, ML models might not be suitable to replace every step in the treatment planning workflow, but they are excellent tools to ease and automate most manual steps.

Eleven dose prediction models were used in our study to increase the number of testing patients instead of a single model. Therefore, a bias towards each model for each subgroup of five testing patients could be derived from this procedure. Nevertheless, the difference between models in the three left-out patients seem to be insignificant in the target volumes and most of the OAR and in any case not higher than 7 Gy for any structure (Appendix C, Figure C.2). In some cases, it might be that for some OAR (e.g the PCMs) the differences shown with respect to the ground truth were also influenced by the prediction model.

It is also important to mention the wide range of applications for automatic RT planning. First, RT plans could be inferred in a few minutes for different treatment modalities (e.g., IMPT versus conventional RT), in order to refer the patient to the most optimal treatment [255, 248, 268]. Second, automatically generated plans could be also used to reduce inter-planner and inter-center variability. Moreover, this approach could be used to inform the planner about where there is room for improvement [269]. Third, the planning workflow which can take several hours or even days, could maybe take an automatically generated plan as a starting point and apply fine-tuning if necessary. Finally, automatic planning could allow for

wide implementation of online adaptive IMPT on high-quality daily images (e.g. repeated CT images or corrected CBCT images [140, 270, 271]). In all previously mentioned scenarios, automatic planning allows for huge time savings, efficient resource usage, or improved plan quality.

In conclusion, this study highlights the crucial necessity of finding the appropriate combination of post-processing settings and dose mimicking strategy to achieve robust deliverable IMPT doses when using automatic planning in oropharyngeal cancer. We stress the importance of a comprehensive robust evaluation of the final dose calculations in the feedback loop to configure a KBP pipeline with the desired clinical specifications.

Chapter 5

Dose mimicking for online adaptation of head and neck cancer

The content is based on a manuscript in preparation with the title "*Dose mimicking based strategies for online adaptive proton therapy of head and neck cancer*". This work was the result of a stay of two months and a half at the Research department of RaySearch Laboratories under the supervision of Albin Fredriksson.

Authors: Elena Borderias-Villarroel¹, Albin Fredriksson², Xavier Geets ^(1,3) and Edmond Sterpin^(1,4)

¹ UCLouvain, Institut de recherche expérimentale et clinique, Molecular Imaging and Radiation Oncology Laboratory, Brussels, Belgium. ² RaySearch Laboratories, Stockholm, Sweden. ³ Cliniques Universitaires Saint Luc, Brussels, Belgium. ⁴ KULeuven, Department of Oncology, Laboratory of external radiotherapy, Leuven, Belgium.

Summary

Purpose: This study compares a not adapted (NA) robust planning strategy with three fully automated online adaptive proton therapy (OAPT) workflows based on the same optimization method: dose mimicking (DM). The added clinical value and limitations of the three proposed OAPT methods are investigated for head and neck cancer (HNC) patients.

Methodology: The three OAPT strategies use DM applied to different reference dose distributions in order to compensate for inter-fractional anatomical changes by re-optimizing new dose distributions on corrected cone beam CT images (corrCBCTs). Order by complexity the OAPTs were: (1) online adaptation based on dose restoration (OADR) where the approved clinical dose on the planning-CT (pCT) was mimicked, (2) online adaptation using DM of the deformed clinical dose from the pCT to corrCBCTs (OADEF), and (3) online adaptation applying DM to a ML predicted dose on corrCBCTs (OAML). Adaptation was only applied in fractions where the target coverage criteria was not met ($D_{98\%} < 95\%$ of the prescribed dose). For 10 HNC patients, the accumulated dose distributions over the 35 fractions were calculated for NA, OADR, OADEF and OAML strategies. DVH metrics and the adaptation rate over the treatment were investigated.

Results: Higher target coverage was observed for all OAPT strategies compared to no adaptation. OADEF and OAML outperformed both NA and OADR and were comparable in terms of target coverage to initial clinical plans. All OAPT methods were able to correct for hotspots. However, only OAML provided comparable NTCP values to those from the clinical dose without statistically significant differences. When evaluating the NA initial plan on corrBCTs, 51% of fractions needed adaptation. The adaption rate was reduced to 25% when selecting the last adapted plan with any of the OAPT strategies as the scheduled plan. The reduction was even more important (down to 12%) when the schedule plan was the best plan over all previously delivered ones.

Conclusion: The implementation of the three OAPT strategies provided superior target coverage compared to no adaptation, higher OAR sparing and reduced the number of needed adaptations.

5.1 Introduction

Intensity-modulated proton therapy (IMPT) offers excellent dose conformity and healthy tissue sparing. Its application is especially interesting in head and neck cancer (HNC) where an important amount of organs at risk (OAR) are in the close vicinity of the tumor. Overdosage to these organs can cause radiation-induced side effects, in particular xerostomia and dysphagia, which negatively impact the quality of life [13, 49]. Unfortunately, IMPT is highly sensitive to uncertainties in patient setup, CT values, and patient anatomy due to displacements of the steep Bragg peak fall-off. Fraction after fraction, changes along the beam path can alter the initially planned Bragg peak causing dose distortions in the approved clinical plan and compromising treatment plan quality [17, 235, 15].

In current clinical practice, robust optimization (RO) for IMPT planning is generally preferred to planning target volume (PTV) optimization. Standard RO usually includes pre-defined setup and range errors during spot weight optimization to generate dose distributions resilient to these specific uncertainties. However, changes in patient anatomy during treatment are harder to address. More sophisticated approaches for RO, such as anatomical RO, explore the incorporation of anatomical scenarios during plan optimization [124, 125, 151]. Nevertheless, it remains difficult to keep the optimality of an IMPT plan while including a wide range of possible anatomical deformations [272].

Although RO can provide a certain level of robustness, it is not the ultimate solution, and treatment plan adaptation is still required in many cases [17, 235, 229, 200]. Usually, offline adaptation triggered by weekly control repeated-CTs (rCTs) or visual inspection of daily CBCTs is used to accommodate for anatomical changes. During offline adaptation recontouring, replanning, treatment plan approval, and additional quality assurance are required in a clinical offline adaptation workflow. As a result, offline adaptation can take several days. In order not to delay treatment delivery and to avoid new anatomical changes during offline replanning, another possibility is to perform online plan adaptation while the patient is on the treatment couch. Since the complete adaptation workflow must be completed in a very short time, online adaptive proton therapy (OAPT) is extremely challenging but it is ideally suited to deal with inter-fractional anatomical changes.

Anatomical changes could be divided into two different types: (1) density changes and (2) non-rigid anatomical changes. A good example of density change in HNC is cavity filling. The anatomy of the patient remains mainly in the same position but it experiences residual rigid translations with an increased density

inside the cavity (from air to mucus). On the other hand, non-rigid anatomical changes imply modifications in the relative position between the OAR and the target volumes. This is the case of complex re-organisations of the anatomy or some tumor shrinkages where the free space left by the reduction in volume is replaced by an OAR that enters the prescription dose region.

Previous studies have shown the potential of dose restoration (DR) strategies to mitigate inter-fractional density changes while maintaining target coverage in lung [237] and prostate cancer [238]. DR is an OAPT strategy that aims at keeping the planned dose distribution stable in every fraction by generating a new (restored) plan with updated proton beam energies and spot weights. Unfortunately, when addressing non-rigid anatomical deformations in HNC, DR is not the most appropriate strategy due to the intrinsic limitations of the method [261].

This study compares a not-adapted (NA) robust planning strategy with three fully automated robust OAPT workflows. The three OAPT methods use dose-mimicking (DM) to generate adapted dose distributions on repeated images. Ordered by complexity the three OAPT strategies are: (1) online adaptation based on dose restoration (OADR) where DM is applied to the reference clinical dose on the planning-CT (pCT), (2) online adaptation using DM on the deformed clinical dose from the pCT to the repeated image (OADEF), and (3) online adaptation using a machine learning (ML) predicted dose on the repeated image followed by DM (OAML). The added clinical value and limitations of the three robust OAPT workflows are compared with respect to no adaptation and between methods in simulated IMPT treatments.

5.2 Material and methods

5.2.1 Patient cohort and treatment planning strategy

Ten bilateral oropharyngeal patients treated with photons were replanned for IMPT delivery to run this study. The prescribed dose (D_p) was 70 GyRBE to the high-dose clinical target volume (CTV) and 54.25 GyRBE to the low-dose CTV in 35 fractions, using a constant relative biological effectiveness (RBE) factor of 1.1. The parotids, submandibular glands (SMGs), the pharyngeal constrictor muscles (PCMs), the oral cavity, the spinal cord, and the brainstem were defined as the main OAR. Target and OAR delineations performed by experienced radiotherapy oncologists were available on all pCTs.

Robustly optimized IMPT treatment plans were semi-automatically planned using the ML model RSL-IMPT-OROPHARYNX-7000-SIB (3.0)*. First, a dose prediction model with UNet architecture [259] provided the predicted IMPT dose distribution. Second, post-processing was applied in order to obtain steep dose volume histograms (DVHs) in the target and spare the OAR (i.e., increase and homogenize the dose to the target while reducing the dose to OAR). Third, robust DM optimization using voxel-wise and regular objective functions was performed. Until this point, all the described steps were fully automated in RayStation 11B and one single click was required to run the optimization. Plans were created using Monte Carlo dose calculation and classical worst-case robust optimization (setup error (SE) = 4 mm, range error (RE) = 3%). SE scenarios included patient shifts in the three axes of the space x,y,z ($\pm$ SE, 0, 0), (0, $\pm$ SE, 0), (0, 0, $\pm$ SE) and the nominal scenario (0,0,0). Range robustness was also included to account for uncertainties in the conversion from CT numbers to proton-stopping power ratios. RE scenarios were modeled by scaling the mass density by $\pm$ 3%. The nominal scenario (i.e., 0%) was also included. The combination of seven SE scenarios and three RE scenarios led to an optimization with twenty-one (7×3) error scenarios.

Finally, automatically generated plans were evaluated against the same robustness parameters used during optimization (SE = 4 mm, RE = 3%). Objective weights were manually adjusted and/or extra optimization objectives were added if necessary until target coverage and OAR clinical goals (see Table 5.1.) were satisfied, in nominal and worst-case scenarios.

The same beam configuration was used for all patients: two anterior oblique beams at ($60^\circ, 10^\circ$), and ($120^\circ, 10^\circ$) and two posterior oblique beams at ($240^\circ, 350^\circ$) and ($300^\circ, 350^\circ$) for gantry and couch angles, respectively. For each patient, the resulting plan from this workflow was considered the reference initial treatment plan.

5.2.2 Anatomy of the day and CBCT correction

CBCT images were acquired with a Halcyon machine (Varian Medical Systems, Palo Alto, CA) during the 35 treatment fractions. Halcyon CBCT images have excellent quality [273] and have been validated for photon dose calculation [274]. Unfortunately, it is well known that artifacts in CBCT images cause unreliable relationships between CBCT intensities and electron densities, especially problematic in IMPT planning. Therefore, correction algorithms are required to enable accurate proton dose calculations [275]. To obtain a more accurate representation of the patient's anatomy of the day, corrected CBCT images (corrCBCTs) were generated in RayStation 11B. The CBCT correction iterative algorithm calculates a correction map and a conversion function for the CBCT based on the pCT.

Priority	Region of interest (ROI)	Clinical goal
Target volume clinical goals		
1	Low-risk CTV	D98% >51.54 Gy (95%Dp = 54.25 Gy)
1	High-risk CTV	D98% >66.50 Gy (95% Dp = 70 Gy)
1	High-risk CTV	D0.03 cm ³ >74.90 Gy (107% Dp = 70 Gy)
2	Low-risk CTV	D98% >53.17 Gy (98% Dp = 54.25 Gy)
2	High-risk CTV	D98% >68.60 Gy (98% Dp = 70 Gy)
Organs at risk clinical goals		
1	Body	D0.03 cm ³ >74.90 Gy (107% Dp = 70 Gy)
1	Spinal Cord	D0.03 cm ³ <45 Gy
1	Brain Stem	D0.03 cm ³ <54 Gy
3	PCMSup/PCMMid/PCMIInf	Mean dose <40 Gy
3	Parotids (left and right)	Mean dose <26 Gy
3	SMGs (left and right)	Mean dose <40 Gy
3	Oral cavity	Mean dose <35 Gy
3	Mandible	D0.03cm ³ <70 Gy
3	Glottic area	Mean dose <40 Gy
3	Esophagus	Mean dose <35 Gy
3	Cochlea (left and right)	Mean dose <30 Gy

Table 5.1: List of clinical goals from the planning protocol used in this study to generate the initial treatment plans .D98% and D2% are the doses received at 98% and 2% of the target volume, respectively. Abbreviations: CTV = Clinical Target Volume, PCM = pharyngeal constrictor muscles, Sup = Superior, Mid = Mid = Middle, Inf = Inferior, SMG: submandibular glands

In other words, the correction algorithm removes the low-frequency artefacts and converts CBCT intensities to HU. To compensate for the limited CBCT field of view (FOV), the areas outside the FOV are filled in by deformably mapping intensity values from the pCT. More information about the method of CBCT correction can be found in section 2.8 of the Reference Manual from RayStation 11B.

Two different contours sets were available in all corrCBCTs: rigidly mapped contours and deformably mapped contours from the pCT to the corrCBCT using rigid registration (RIR) and deformable image registration (DIR), respectively. The DIR algorithm was the ANACONDA algorithm available in RayStation 11B [276].

5.2.3 Adaptive strategies

Dose mimicking

The three proposed OAPT strategies used RayStation dose mimicking (RSDM) as the optimization method to generate online adapted plans in the corrCBCTs. The RSDM used in this study corresponds to the mimicking method integrated with the ML model RSL-IMPT-OROPHARYNX-7000-SIB (3.0) [260].

RSDM is formulated as a voxel-wise least squared error (LSE) minimization problem. The LSE objective function is divided into two one-sided terms, enabling different weights depending on whether the optimized dose is greater or less than the reference dose given as input. For OAR, voxels with smaller dose values than the reference dose are not penalized. In addition to the voxel-wise LSE objective, a weighted sum of conventional radiotherapy objective functions is added to the problem, as implemented in RayStation 11B.

RSDM procedure consists of two separate runs, between which the clinical dose is calculated accurately. These two runs consist of 125 and 80 iterations respectively, which in our study were completed in between 13 to 17 minutes for the 21 robust scenarios. Optimizations were run on a GPU0 Quadro P5000, 15 GB, WDDM.

For the moment, proton dose computation on corrCBCTs is not available clinically, but only in research versions. In collaboration with RaySearch, the RSDM model was adapted to run on corrCBCT instead of pCT and to receive as input not only an ML predicted dose but any other given dose distribution. When using RSDM, the contours of the ROIs to which the optimization objectives are applied also need to be specified.

5.2.4 Online adaptation strategies

Three OAPT workflows were fully automated using the RayStation scripting interface. The ROIs and reference doses given as input to the RSDM to run new plan optimizations on the corrCBCTs are shown in Figure 5.1.

- **Online Adaptation using Dose Restoration (OADR).** The initial reference dose and manually delineated contours on the pCT were rigidly mapped to the corrCBCT and given as input to the RSDM.
- **Online Adaptive using Deformable image registration (OADEF)** . The initial reference dose and manually delineated contours on the pCT were mapped to the corrCBCT using DIR and given as input to the RSDM.

- **Online Adaptive using a Machine Learning predicted dose (OAML).**
Manually delineated contours on the pCT were deformably mapped to the corrCBCT. The deformed contours were given as input to the ML model to predict a new dose distribution on the corrCBCT. Later, the RSDM was applied to the predicted dose using the deformed contours.

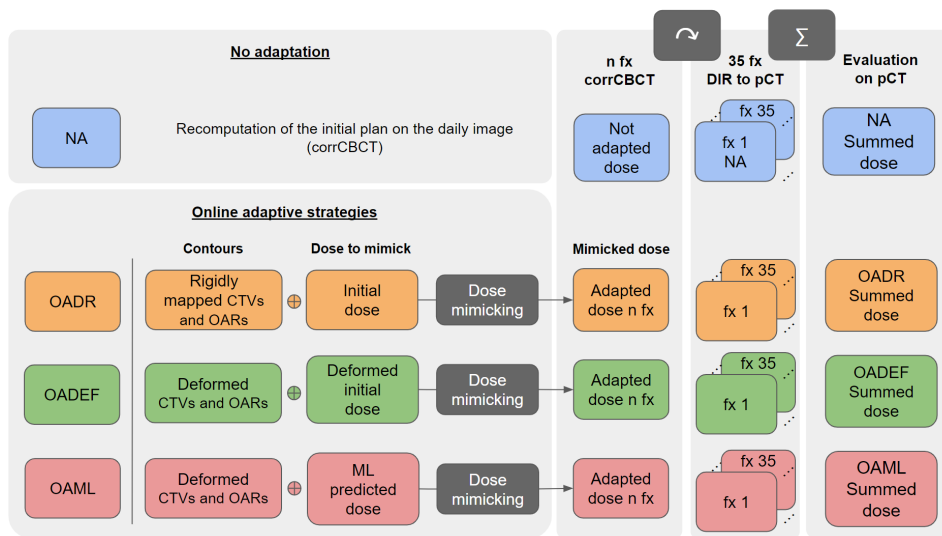


Figure 5.1: Summary of the workflow to compare the not adaptive (NA) strategy with the three proposed online adaptive methods based on dose mimicking: OADR, OADEF and OAML. Abbreviations: corrCBCT = corrected CBCT, DIR = Deformable Image Registration, fx = fraction, ML= Machine Learning, NA = No Adaptation, OADR = Online Adaptive based on Dose Restoration, OADEF = Online Adaptation based on DEformed dose, OAML = Online Adaptation based on ML predicted dose, pCT = planning-CT.

5.2.5 Adaptation schedule

The objective was to simulate online adaptation only in the fractions where it was needed rather than to explore a daily adaptation schedule for each strategy. To trigger the adaptation, initial plans were recomputed on the corrCBCT and the target coverage was evaluated on the propagated (i.e deformed) contours from the pCT to the corrCBCT. A fraction was flagged as requiring adaptation when the target coverage criteria (i.e $D_{98\%CTV} < 95\%$ Prescribed dose (D_p)) was not met for high-risk and/or low-risk deformed CTVs. From the 10 patients, 7 needed adaptation in more than 15 fractions (from 33 to 16), while 3 required only adaptation in 1 to 10 fractions.

5.2.6 Evaluation strategies

For not adapted treatments (NA), the dosimetric effects of anatomical and setup uncertainties were calculated by recomputing the initial plans on the corrCBCTs. The adapted doses in each fraction were simulated by optimizing a new adapted plan on the daily image (corrCBCT), using the different OAPT methods (OADR, OADEF and OAML) as described in Figure 5.1.

With 35 fractions per patient and a total of 10 patients, 350 fractions were available to evaluate target coverage on the deformed CTVs for NA, OADR, OADEF and OAML strategies. Ideally, the evaluation of the OAPT strategies should be performed on manually delineated contours on the corrCBCTs for each fraction. However, the re-delineation of 350 corrCBCTs would imply an unreasonable workload for the clinical staff at our institution. As an alternative, we decided to use dose accumulation on the pCT in order to simulate a complete treatment delivery (see Figure 5.1). A verification on a subset of 2 corrCBCT per patient, corresponding to 20 treatment fractions, was performed on re-contoured targets and OAR.

Evaluation on accumulated doses

According to the initial plan adaptation schedule (section 5.2.5) for each patient, NA, OADR, OADEF and OAML plans were computed on the 35 corrCBCTs to simulate the final delivered dose by each method. Using DIR (from the pCT to corrCBCT), each fraction dose was warped back to the pCT. The 35 warped fraction doses were summed up to four different accumulated dose distributions per patient (NA, OADR, OADEF and OAML summed doses), and compared to the initial reference dose. It is important to keep in mind that daily adaptation was not performed. Therefore, OADR, OADEF and OAML summed doses contained NA doses in the fractions where adaptation was not required (i.e target coverage criteria were already met by using the initial clinical plan without adaptation for that specific treatment fraction). Since all the doses were warped back to the pCT, the metric evaluation was performed on the manually delineated contours on the pCT.

For each summed dose, differences with respect to the clinical dose in D98 values (dose received by 98% of the volume) were calculated for low-dose and high-dose CTVs. Differences in Dmean (average dose) were reported for the parotids, SMGs, PCMs and the oral cavity while differences in maximum doses (i.e dose received at 0.03 centimeters cube (D0.03cc)) were calculated for the spinal cord, brainstem and the body. These metrics were evaluated for NA, OADR, OADEF and OAML accumulated doses and compared to the initial dose distribution values.

Median, minimum and maximum values of the differences, Median (Min, Max), were reported to illustrate how far the NA and OAPT doses were from the initially planned dose distribution. Normal Tissue Complication Probability (NTCP) values were also calculated from the OAR metrics using the formulation proposed by Langendijk et al.[277].

Evaluation on corrected contours in corrCBCT

The dose accumulation method previously described has limitations since the DIR used to deform the contours from the pCT to the corrCBCT was also used (inversely) to map the NA, OADR, OADEF and OAML dose distributions in each corrCBCT to the pCT. It could be that inaccuracies in the DIR used for contour mapping would have been underestimated because of the posterior inverse dose mapping to the pCT. In this case, OADEF and OAML strategies would have biased evaluations giving too optimistic results.

To overcome the DIR limitations we asked experienced radiation oncologists to correct CTV and OAR volumes in two fractions per patient. Fractions with low target coverage in the deformed CTVs were selected: one fraction before the first half of the treatment and another one towards the end of the treatment. CTV and OAR contours were deformably mapped and manually corrected when necessary. With two fractions for each of the 10 patients, 20 fractions with manually corrected contours were available for direct evaluation of each OAPT strategy.

Evaluation of different treatment schedules

The scheduled plan is defined as the plan that should be delivered to a patient at a specific n treatment fraction ($n \leq 35$). In this study, we evaluated for each of the OAPT strategies the number of fractions needing adaptation depending on different scheduled plans for each fraction n :

- **Initial plan (IP):** The IP optimized on the pCT was the scheduled plan for fraction n .
- **Last adapted plan (LAP):** From all the previous fractions, the plan corresponding to the last adapted fraction X is the scheduled plan for fraction n ($X < n$).
- **Best adapted plan (BAP):** The plan with the highest target coverage after the re-computation of all previously adapted plans on the corrCBCT of the fraction n is chosen as the BAP. The BAP corresponding to a previous fraction Y ($Y \leq X < n$) is the scheduled plan for fraction n .

5.2.7 Statistical significance

Grouped data were tested for statistical significance using two-sided Wilcoxon-Mann-Whitney tests using SciPy version 1.5.2. The differences were considered statistically significant for p -values < 0.05 .

5.3 Results

5.3.1 Target coverage in corrected CBCTs

Median values of target coverage across the 350 corrCBCT (including adapted and not adapted fractions) are given as a percentage of the prescribed dose in Table 5.2.a) for low and high-risk CTV. With NA strategy relying on the robustness of the initial plan, 52%(181/350) of the fractions required adaptation since target coverage criteria ($D98\% > 95\%D_p$) was not fulfilled for one or both CTVs. When applying OADEF or OAML to these fractions, target coverage was ensured in all corrCBCT. This was not the case with OADR, which could not ensure target coverage in 9% and 5% of the cases for low-risk and high-risk CTV, respectively. For all strategies, NA, OADR, OADEF and OAML coverage of low-risk CTV was lower than high-risk CTV. OADR improved median target coverage in 0.9 Gy compared to NA plans. While OADEF had the highest performance with less than 0.2 Gy difference with respect to OAML.

5.3.2 Target coverage in accumulated doses

Table 5.2.b contains the median values of D98 metrics for both CTVs across the 10 accumulated dose distributions (one per patient). Without any adaptation, low-risk CTV did not receive the prescribed dose in 3 patients (30%). When implementing OADEF and OAML to the selected fractions for adaptation, accumulated dose distributions recovered target coverage in all patients. For OADR, the low-risk CTV of patient 3 received only 93.46% of the prescribed dose, being the only patient for which the method was not able to compensate for the anatomical changes.

Differences in D98% values of NA, OADR, OADEF and OAML strategies with respect to the initial plan are shown in Figure 5.2 Whether the differences with respect to the initial dose and between models are statistically significant ($p < 0.05$) or not statistically significant (n.s, $p > 0.05$) was reported in Table 5.3.

NA plans presented statistically significant degradation in target coverage with respect to the initial plan. In the high-risk CTV the median value of the difference for D98 metrics NA plans was -0.84 (-1.21, 0.11) Gy Gy. On the contrary, when

a) Evaluation on deformed CTVs for the 350 corrCBCTs				
Median (Min - Max) % of prescribed dose				
Planning strategy	Low risk CTV (Dp = 54.25 Gy)	Fail cases Abs (%)	High risk CTV (Dp = 70.00 Gy)	Fail cases Abs (%)
NA initial plan	95.21 (60.77 - 96.79)	151 (43%)	96.11 (73.38 - 97.89)	117 (33%)
OADR	96.13 (74.35 - 97.24)	34 (9%)	97.04 (84.39 - 98.20)	18 (5%)
OADEF	96.54 (95.01 - 97.43)	0 (0%)	97.52 (95.08 - 98.22)	0 (0%)
OAML	96.43 (95.01 - 97.27)	0 (0%)	97.38 (95.08 - 98.10)	0 (0%)
b) Evaluation of the 10 accumulated dose distributions on the planning contours				
Median (Min - Max) % of prescribed dose				
Planning strategy	Low risk CTV (Dp = 54.25 Gy)	Fail cases Abs (%)	High risk CTV (Dp = 70.00 Gy)	Fail cases Abs (%)
NA initial plan	96.07 (92.73 - 97.42)	3 (30%)	96.47 (96.20 - 97.89)	0 (0%)
OADR	97.30 (93.46 - 98.01)	1 (10%)	97.61 (96.58 - 98.14)	0 (0%)
OADEF	97.87 (97.04 - 98.24)	0 (0%)	98.03 (96.52 - 98.84)	0 (0%)
OAML	97.82 (97.10 - 98.07)	0 (0%)	97.90 (97.07 - 98.57)	0 (0%)
c) Evaluation on manually delineated CTVs for 20 selected corrected CBCTs				
Median (Min - Max) % of prescribed dose				
Planning strategy	Low risk CTV (Dp = 54.25 Gy)	Fail cases Abs (%)	High risk CTV (Dp = 70.00 Gy)	Fail cases Abs (%)
NA initial plan	93.74 (59.76 - 95.99)	16 (80%)	95.05 (89.13 - 97.52)	9 (45%)
OADR	95.83 (90.18 - 96.94)	4 (20%)	97.36 (86.66 - 98.10)	3 (15%)
OADEF	96.70 (95.96 - 96.91)	0 (0%)	97.80 (92.54 - 98.19)	1 (5%)
OAML	96.50 (96.12 - 96.89)	0 (0%)	97.66 (90.48 - 98.04)	1 (5%)

Table 5.2: Median, minimum and maximum values for low-risk CTV and high-risk CTV target coverage for three different evaluation strategies: (a) evaluation on deformed CTVs over the 350 fractions (10 patients with 35 fractions per patient), (b) evaluation on the manually delineated contours on the planning-CT after dose accumulation of the 35 fractions on the planning-CT (one dose distribution per patient, so 10 accumulated dose by strategy) and (c) evaluation on the corrected CTVs on a subset of 20 selected fractions. The number of cases where the target coverage criteria failed ($D_{98} < 95\% D_{prescribed}$) is also reported in absolute value (Abs) and as a percentage of the number of total evaluated data points: 350 for (a), 10 for (b) and 20 for (c). Abbreviations: Abs = Absolute, CTV = Clinical Target Coverage, NA = Not adapted, Min = Minimum, Max = Maximum.

applying OAPT strategies median differences were not statistically significant with respect to the initial plan with 0.01 (-1.17, 0.62) Gy for OADR, 0.29 (-1.22, 0.94) Gy for OADEF and 0.16 (-0.83, 0.61) Gy for OAML.

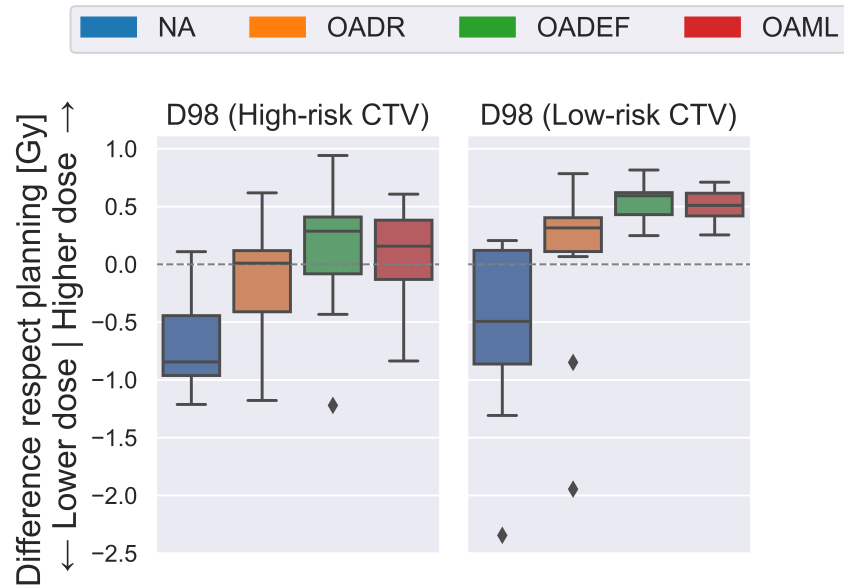


Figure 5.2: The differences in terms of target coverage (D98% - dose received by 98% of the CTV) between accumulated doses from OAPT strategies and the clinical plans. Negative differences indicate an underdosage of the CTV with respect to the clinical plan while higher target coverage is linked to positive differences. The boxes indicate the median and interquartile range (IQR). The whiskers extend to a minimum of 1.5 times the IQR and diamonds represent the outliers.

For the low-risk CTV (i.e union between both left and right nodes) differences in D98 values experience a drop in target coverage of -0.50 (-2.34, 0.20) Gy for NA plans. On the contrary, when applying OAPT strategies, low-risk target coverage was increased even to greater values than the ones from the reference plan, with a median D98 values of 0.32 (-1.95, 0.78) Gy for OADR, 0.59 (0.25, 0.82) for OADEF and 0.51(0.25, 0.71) for OAML. For OAML and OADEF the differences with respect to the clinical dose were statistically significant.

	D98(Low risk CTV)				D98(High risk CTV)			
	NA	OADR	OADEF	OAML	NA	OADR	OADEF	OAML
Clin	0.008	n.s	0.002	0.002	0.004	n.s	n.s	n.s
NA		0.010	0.002	0.002		0.019	0.006	0.002
OADR			0.003	n.s			0.019	0.027
OADEF				n.s				n.s

Table 5.3: Statistical significance of the different strategies (NA, OADR, OADEF and OAML) compared to the reference clinical dose and between them. Statistically significant (s.s) implies $p < 0.05$. Pink cells represent the values for which no statistically significant difference (n.s.s) was reported ($p > 0.05$).

Analysing the performance between methods, OADEF provides the highest target coverage overall but with an outlier for patient 10 for which only one fraction was adapted. OADEF and OAML perform very equally, ensuring target coverage for all accumulated dose distribution. The differences in target coverage with respect to the initial dose for both low-risk and high-risk CTV between OADEF and OAML were not statistically significant.

5.3.3 Target coverage for re-delineated contours

Median values of target coverage across the 20 re-delineated corrCBCT were reported as a percentage of the prescribed dose in Table 5.2.c) for both CTVs. When the initial plan was delivered without any correction, adaptation was required in 80% and 45% of the 20 fractions to ensure target coverage for low-risk and high-risk CTV, respectively. In some cases, the target coverage was extremely low, not even achieving 60% of the prescribed dose for one outlier. With no adaptation, the median value of D98% metric for the low-risk CTV was below 94% of the prescribed dose. When applying OADR to this subset of fractions, median D98% was recovered above 95% of the prescribed dose for low-risk CTV and increased from 95% to 97% of the prescribed dose for high-risk CTV. However, some of the fractions still did not reach the clinical goal ($D98\% > 95\% D_p$). Although OADEF and OAML strategies increased target coverage for both CTVs, none of the methods could ensure coverage of the high-risk CTV for one of the fractions when evaluating the real (i.e corrected) structures.

5.3.4 OAR sparing and toxicity on accumulated doses

The differences with respect to the initial dose in the OAR and NTCP evaluation metrics were displayed in Figure 5.3 and Figure 5.4, respectively. Median values of the differences in the OAR metrics were in most of the cases less than 1 Gy. A bigger variability was shown in the PCMs, especially on the superior and medium PCMs. Overdosage to the PCMs was translated into a higher probability of dysphagia in grades 2 and 3, especially for NA plans. When applying OADR and OAEF, NTCP values were reduced close to the initial values. The lowest NTCP values were achieved by OAML doses, with no statistical significance compared to NTCP values from the initial plans.

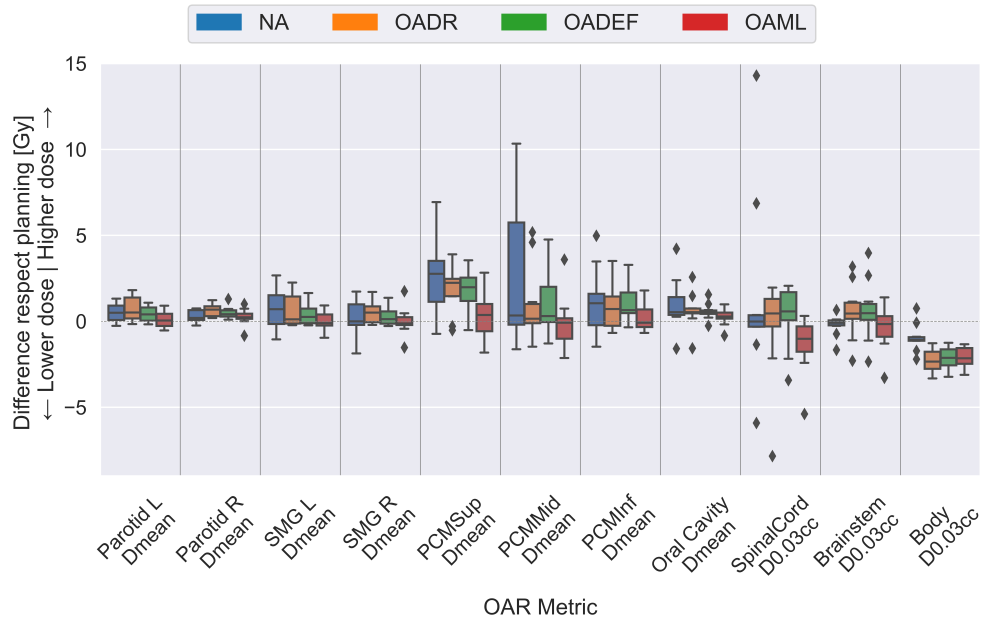


Figure 5.3: The differences in DVH metrics (Dmean and D0.03cc) between accumulated doses of NA, OADR, OAEF and OAML, and the clinical plans. Differences greater than zero indicate an overdosage to the OAR while values below zero indicate OAR sparing. The boxes indicate the median and interquartile range (IQR). The whiskers extend to a minimum of 1.5 times the IQR and diamonds represent the outliers.

Two outliers were observed in the spinal cord for NA plans with overdoses of 6.8 Gy and 14.5 Gy. Both were compensated with the implementation of any of the OAPT strategies. Maintaining target coverage came at the cost of a slightly increased dose to the brainstem and the spinal cord for OADR and OADEF strategies, although always within clinical limits. Moreover, an increased dose to the parotids in OADR dose distributions was translated into a median increase of 0.5% probability for xerostomia grade 2. The NTCP values never exceeded more than 1.2% for xerostomia nor 3.9% for dysphagia. Moreover, in the case of dysphagia, NTCP values were always lower when applying any of the OAPT strategies than with no adaptation.

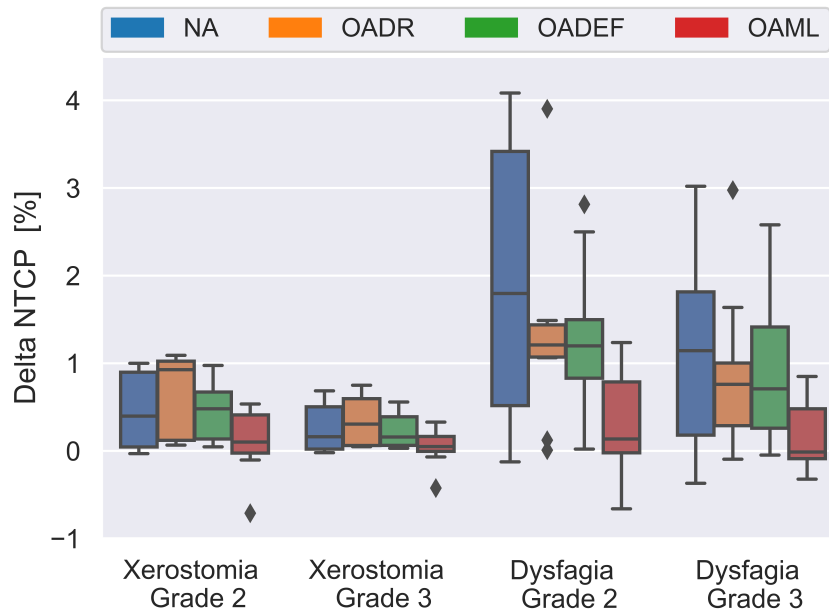


Figure 5.4: Difference in NTCP values respect to the clinical plan NTCP values, $\Delta\text{NTCP} = \text{NTCP}(\text{OAPT strategy}) - \text{NTCP}(\text{clinical plan})$. The higher the difference in the NTCP values, the bigger the deviation from NTCP from the clinical plan and the higher probability of developing secondary effects associated with toxicity (for values above 0).

5.3.5 Analysis of adaptation rates

In Table 5.4. the adaptation rates for three different scheduled plans: initial plan (IP), last adapted plan (LAP) and best adapted plan (BAP) were reported. More than half of the fractions (51%) required adaptation to ensure target coverage criteria when the initial plan was chosen to be delivered without any adaptation. On the contrary, when LAP was selected as the scheduled plan, the need for adaptation was reduced by more than half for any of the OATP strategies: 25% of adaptation required for OADR, 16% for OADEF and 21% for OAML. The adaptation rate decreased, even more, using the BAP as the scheduled plan. In this case, when applying 20% of adaptation was achieved for OADR, 14% for OAML and in the best case scenario, only 12% of the 350 fractions (instead of 51%) was required when OADEF was chosen.

Patient	Number of adapted fractions						
	Initial plan (IP)	Last adapted plan (LAP)			Best adapted plan (BAP)		
		OADR	OADEF	OAML	OADR	OADEF	OAML
1	33	28	14	21	20	11	14
2	32	13	12	14	9	8	10
3	26	20	7	7	19	5	7
4	22	5	5	5	4	4	4
5	20	5	4	5	5	4	4
6	18	4	3	11	3	3	4
7	16	5	5	5	5	2	2
8	10	5	4	5	3	3	4
9	2	2	2	2	2	2	2
10	1	1	1	1	1	1	1
%Total of the 350 fractions	51%	25%	16%	21%	20%	12%	14%

Table 5.4: The number of adapted fractions per patient according to the target coverage criteria ($D_{98\%} > 95\%$ Prescribed dose) evaluated in each fraction. Since manually delineated contours were not available in every fraction, the recomputation of the schedule plans for that fraction (i.e the initial plan (IP), the last adapted plan (LAP) or the best adapted plan (BAP)) was evaluated on the deformed target volumes.

5.4 Discussion

Three robust IMPT planning strategies based on dose mimicking were fully automated and integrated into the commercial TPS RayStation 11B. The three methods, OADR, OADEF and OAML, increased target coverage compared to no adaptation while reducing the need for offline adaptation. It was shown that the adaptation rate could be reduced from 51% for NA plans, to less than 25% when using LAP as the scheduled plan, and below 12% for BAP. BAP is computationally more expensive than LAP since it requires more than one proton dose recalculation. However, the workflow could be easily parallelized to quickly evaluate all previously adapted plans.

Due to the proximity of small OAR to the CTVs in HNC, maintaining the target coverage during all fractions slightly increased the dose to the parotids and PCMs. This was translated into increased NTCP with respect to the initial plan, especially for OADR and OADEF strategies. However, for dysphagia, the increase in NTCPs was always lower for any OAPT strategies than for NA, while for xerostomia it never exceeded more than 2%. This could seem an acceptable tradeoff in exchange for a higher target coverage and tumor control probability (TCP). Overall, OAML dose distributions showed the best performance, maintaining or even improving target coverage and OAR sparing and with no statistical significance in NTCP values with respect to the initial plan for all toxicities and grades. Moreover, all three OAPT strategies were also able to correct hot spots.

Previously presented robust dose restoration strategies required an additional manual step in the planning phase for each patient [237, 261, 33]. This approach could make difficult the usability of the method in a clinical setting due to the increased workload [146]. On the contrary, in our proposed OADR, the manual configuration of the mimicking parameters is performed once during the commissioning of the ML model. Once the mimicking settings are chosen for a specific protocol (i.e tumor location, beam configuration, robustness parameters etc.), the mimicking model can be used for all subsequent patients without manual tuning.

Manual contour delineation of the new patient anatomy is the most time-consuming step in the adaptive workflow and probably one of the bottlenecks hampering online adaptive implementation [278]. The advantage of the three proposed OAPT strategies is that they do not require new contour delineation before re-optimization, contours from the pCT are instead (a) rigidly or (b) deformably mapped to the corrCBCT.

When (a) rigid mapping is applied, we gain stability in the dose shape compared to the initial clinical plan, which facilitates plan approval and quality assurance (QA). Only addressing density changes, the offline adaptation rate was not zero but it was reduced by more than 50%. To cover a wider range of anatomical deformations, including non-rigid deformations, other adaptive strategies including deformed contours (b) were required. These results are in agreement with a previous study by Borderias-Villarroel et al. [261] where offline adaptation was necessary to correct the dose distortions that dose restoration could not compensate for. The proposed OADEF and OAML strategies could be alternatives to offline replanning in such cases to avoid increased workload and treatment delays. Ideally, the re-contouring (i.e volume segmentation) could be performed automatically by artificial intelligence models [242, 246]. These tools have been proven very effective on CT image quality data sets but would need to be adapted to be used on low-quality CBCT images [279, 280] or synthetic CTs [271]. Unfortunately, nowadays both DIR and automatic segmentation cannot be relied on blindly, without human verification [236]. Therefore, contour modification and approval by the doctor are still required, as is the case in the Ethos System from Varian or the MRLinacs [194, 136].

Although the benefit of adaptation is already seen with only some of the fractions being adapted, it could be that its main advantages, such as higher OAR sparing, are diluted by the non-adapted fractions. A bigger difference between NA plans and OAPT plans would have probably arisen from a comparison with a daily adaptation schedule. However, with the current technology daily online adaptation is still a cumbersome task that will increase the workload and the time spent by the patient on the treatment couch. This could be detrimental for the patient due to increased setup error uncertainties in fractions where adaptation is not needed to achieve the clinical goals. On the other hand, higher adaptation rates open the path to decreasing robustness parameters accounting for setup and range errors that could be translated into OAR sparing and reduced toxicities.

The feasibility of fully integrated and automated adapted proton plans using Monte Carlo optimization on corrected CBCT images in less than 20 minutes, brings OAPT one step closer to its clinical implementation. However, some of the limitations of our study still need to be addressed.

First, the CBCT image quality was out of the scope of the study since our main objective was the comparison between adaptive strategies all computed on the same corrCBCT. Nevertheless, it is an important field of research in OAPT since an improved CBCT image quality would enable more accurate evaluation, optimization and proton dose calculation [140, 281, 171].

Second, the DIR algorithm introduces uncertainties during contour mapping, dose warping and accumulation especially when important anatomical changes occur and when reporting small volume dose metrics (i.e parotids or SMGs) in steep dose gradients [282, 283, 158]. In our study, NTCPs were calculated with OAR median doses taken directly from the DIR-based deformed contours. However, differences in NTCPs between DIR-based and manual contouring should be within the range caused by inter-observer variability as reported by Gan et al. [284]. It was also shown the evaluation on deformed contours can be suboptimal to estimate the real delivered dose to the patient leading to an underestimation of the adaptation need. The evaluation of OAPT strategies directly on a subset of fractions with manually corrected contours attempted to mitigate the limitations of the DIR-based dose accumulation. However, this probably came intrinsically with the introduction of inter-observer variabilities. Moreover, with the introduction of DIR-mapped contours in automatic proton treatment re-planning, there is a crucial necessity to develop automatic tools able to identify when propagated contours are likely to be incorrect, acting as a flag for manual review of these contours [285].

Third, depending on the available tools and medical staff OADR, OADEF, OAML or offline adaptation could be chosen as adaptive strategies. A fast-independent Monte Carlo dose calculation could help in the decision-making and as a pre-treatment patient-specific QA procedure [286]. Parameters such as the degree of weight loss, reduction in lateral neck diameter or Dice Similarity Coefficient (DSC) between rigidly and deformably mapped contours could also be monitored to trigger patients receiving the greatest benefit of adaptation and appropriate candidates for each of the OAPT [261, 243]. In any case, all online adapted plans should be followed by an extensive and fast, patient-specific quality assurance including machine log files after delivery [104]. Finally, it would be interesting to evaluate the three OAPT methods compared to reference clinical plans manually generated instead of fine-tuned automated ML plans.

In conclusion, all adapted plans provided closer DVH metrics to the validated initial plan compared to no adaptation, sometimes reestablishing or even improving the initial plan. OAML had the overall better performance in target coverage, OAR sparing and NTCPs, closely followed by OADEF. The three methods have the potential to be integrated into an online adaptive clinical setting to generate adapted plans which can compensate for inter-fractional anatomical changes and reduce the need for offline adaptation. This is key to deliver the most personalised treatment plan to the patient while avoiding an unaffordable increased workload to the radiation therapy department.

Chapter 6

Conclusion and outlook

Chapter 6 concludes this work and presents the outlook and perspectives for future work.

The Bragg peak is the main advantage of proton therapy, but it is also its main challenge. In contrast to the low sensitivity of photon treatments, proton dose distributions are distorted in the presence of treatment uncertainties due to range errors. Online APT can be implemented to compensate for inter-fractional anatomical changes and realize the full potential of proton therapy. This thesis aims at this objective by exploring fully automated robust planning strategies in a commercial TPS. The contributions from the author can be summarized as:

- An efficient method for fully automated online APT based on dose restoration was successfully implemented and validated in several patients with inter-fractional anatomical changes. First, dose restoration was investigated for lung cancer patients, including robustness against breathing motion (Chapter 2). Later, the dose restoration workflow was modified to be applied to HNC patients (Chapter 3). Dose restoration managed to reconstruct the prior approved dose distributions as made for the pCT scans despite density changes on rCTs. Thus, the aim of replicating the static dose cloud approximation valid for photons was also achieved in proton treatments by compensating for proton range errors. The need for offline plan adaptation was reduced when implementing the dose restoration adaptive approach.
- The feasibility of fully automated IMPT plans combining post-processed predicted doses and dose mimicking optimization was investigated in Chapter 4. A particular focus was put on the post-processing and mimicking steps since they are essential to ensure the accuracy and robustness of the automatically generated plans. From comparing four different variations of the planning workflow, information could be extracted to derive the appropriate combination of post-processing and dose mimicking settings that provided the most accurate and robust IMPT plans.
- Three automatic IMPT planning strategies using dose mimicking were developed to be applied on HNC repeated CBCTs (Chapter 5). Each strategy received a different dose to mimic to generate an adapted dose: the initial clinical dose (OADR), the deformed clinical dose (OADEF) or a completely new reference dose generated via machine learning dose prediction (OAML). Although the proposed strategies differ in the requirements for their clinical implementation and their range of application, all demonstrated their potential to perform successful treatment plan adaptation in different anatomical changes scenarios. These automatic planning methods could be used in an online adaptive setting to generate adapted IMPT plans to the anatomy-of-the-day.

The proposed methods may benefit cancer patients through more accurate, robust, and patient-specific proton therapy treatment delivery using online APT. More accurate since online APT has the potential to ensure that what was initially planned is indeed delivered without delays during the treatment course. More robust because treatment uncertainties can be taken into account during re-optimization of the adapted plans. Finally, more patient-specific due to the possibility of daily treatment plan adaptation to compensate for patient anatomical changes at every treatment fraction. However, there is still room for improving the presented methods, as well as there are other related topics not treated here that should be further investigated. In this chapter, we will review the remaining issues and limitations of the developed methods, followed by a discussion of the opportunities and challenges in their clinical implementation and the future of online APT.

6.1 Potentials and limitations of the proposed online APT methods

Dosimetric improvements

In agreement with the proof of concept of isodose-based dose restoration in two HNC patients and one lung cancer patient by Bernatowicz et al. [33], we also found improved DVH metrics of the restored plans with respect to no adaptation for larger patient cohorts of lung and HNC. As discussed in Chapter 2 for lung cancer patients, target coverage and OAR sparing were always improved (i.e., closer to initial planning values) when applying dose restoration compared to no adaptation. However, dose restoration was not that effective in the investigated HNC cohort (Chapter 3). Although dose restoration yielded adequate target coverage in 52% of rCTs, full offline adaptation was still required in 48% of rCTs to achieve acceptable dose levels. Changes in the relative positions of anatomic structures due to non-rigid deformations or poor quality registrations caused imperfect overlap between rigidly mapped and the real position of the regions of interest (CTV and OAR). In such cases, although dose restoration successfully replicated the initial dose distribution on the rCT, the dose was restored in the wrong location. In the presence of such non-rigid anatomical changes, dose restoration could even be detrimental compared to no adaptation. Therefore, it is crucial to (1) identify such situations and (2) to find more appropriate adaptive strategies to suit them.

To address (1), the Dice Similarity Coefficient (DSC) between rigidly and deformably mapped target volumes or other anatomical indicators (e.g., neck diameter) were proposed in Chapter 3. Nevertheless, further research is required

to define a mechanism capable of identifying patients for whom dose restoration would be inadequate. Alternatives using deformably mapped contours and dose mimicking for automatic IMPT planning were investigated in Chapter 5(2) to overcome the intrinsic limitations of the dose restoration approach. These online APT strategies using deformed contours (e.g OADEF and OAML) provided superior dosimetric scores than the restored and not adapted plans for both target coverage and OAR sparing. Nevertheless, the use of contour deformation comes at the cost of contour verification, requiring the presence of the doctor in the treatment room. This point will be discussed further in this section.

Offline adaptation rate reduction

Nowadays, proton clinics rely on rCTs to trigger offline adaptation to handle anatomical changes. Evans et al. highlighted the importance of weekly verification rCTs for HNC since they frequently influenced clinical decision-making to perform re-planning ($\sim 50\%$) [25]. As previously mentioned, offline adaptation implies delays in the treatment delivery to the patient and an increased workload in the radiation therapy department. Launching the complete offline APT workflow from scratch can take several days. Although some automation efforts based on DIR have achieved timing reductions below 24 hours [287], the more this cumbersome task can be reduced, the better. The better for the patient who will receive the treatment initially planned without delays, and the better for the radiation therapy service, which can spare time and human resources.

When analyzing the offline adaptation rate based on target coverage, it was demonstrated in Chapters 3 and 5 that dose restoration techniques could account for a high number of dose distortions reducing the need for offline APT. Only one out of 14 patients required offline adaptation in the lung patient cohort. In the investigated HNC cohort, dose restoration reduced the need for offline adaptation in 50% of the cases. However, although dose restoration could reduce offline adaptation rates to numbers more typical of conventional photon therapy, it is not the ultimate solution for DAPT since it can not properly handle some non-rigid anatomical changes. Therefore, once the proton-specific problem of range errors has been mainly solved by dose restoration, the residual adaptation rate associated with wide non-rigid anatomical changes is a challenge shared between photon and proton radiation therapy.

In Chapter 5 we saw that the residual need for adaptation could be handled by the strategies using non-rigidly mapped contours (i.e OADEF and OAML). OADEF and OAML did not need offline APT when evaluating accumulated doses in the investigated HNC patient cohort. The only exception was a single rCT (out of 20) in which clinical goals for target coverage were not achieved in manually

delineated contours. These results highlighted the ability of these methods to properly cope with dose distortions induced by a wide range of inter-fractional anatomical changes. Moreover, they opened the path to replace, or at least minimize, the need for offline adaptation. Finally, they contribute to the previous literature [198, 288, 238] to fill the gap towards the clinical implementation of a DAPT workflow, which is of growing interest for proton therapy [135].

Accuracy, ensuring target coverage

While developing the adaptation methods described in this thesis, the focus was on making the workflow fully automated and free from human intervention. To this end, we aimed for a success rate of 100% in the adapted plans (i.e., meeting all clinical goals). This entails making the adapted plans clinically acceptable without user alterations, thereby reducing the need for review. Therefore, strict criteria for target coverage were applied while lower priority was given to the OAR sparing. The dose restoration method achieved this aim using higher optimization weights in the isodoses corresponding to the dose prescription region. For dose mimicking strategies, ensuring target coverage relied on the post-processing of the mimicked dose and high optimization weights on target structures during dose mimicking process.

However, slightly lower target coverage could be clinically desirable in some instances, as critical OAR could be better spared. If this holds for the entire treatment course, one could generate several models based on dose mimicking with different post-processing and mimicking settings that would push to spare specific OAR. Depending on the treatment fraction, different models could be chosen by the medical doctor when triggering the adaptation. However, as shown in Chapter 4, this could highly compromise the robustness of the plans. This approach should be, therefore, further investigated for safe implementation. Following the same reasoning for the isodose-based dose restoration method, additional patient-specific templates could be created offline at the planning stage to spare specific OAR. However, the additional workload that this would imply does not seem reasonable in a clinical setting.

6.2 Application of online APT strategies

Potential for margin reduction

We have seen that treatment uncertainties are often handled using margins, robust optimization, and treatment plan adaptation. The more uncertainties we consider upfront during the initial treatment planning, the more robust our plan

will be, and less adaptation will be required. However, the higher robustness of the initial plan comes at the cost of a larger irradiated volume which can undermine the potential benefits of proton therapy. As shown by Van de Water et al. [289], reducing the amount of robustness, especially for the setup error, could reduce NTCP in oropharyngeal cancer patients.

Ideally, with the clinical implementation of DAPT, the degree of included robustness/margins could be further reduced as the initial plan no longer has to account for all possible shapes and locations of the target. Furthermore, this reduction could be translated into an even more conformal proton dose distribution with a higher sparing of the healthy tissue and, therefore, into reduced toxicity. Nenoff et al. [198] demonstrated that DAPT mitigated daily positioning errors, allowing for a substantial reduction of PTV margins, thus decreasing integral dose while preserving target coverage throughout the treatment. Wagenaar et al. also demonstrated the feasibility of a reduced 2 mm setup margin for HNC [290]. However, the clinical implementation of such approaches would benefit from a DAPT strategy that could be safely applied in case of unexpected dose distortions during the treatment course. When combining online APT with other novel techniques, such as dual-energy CT, the robustness setting reduction could be even higher [291].

A recent proof-of-concept study by Dohopolski et al. compared daily ART treatments of 1 mm PTV margin to the conventional 5 mm PTV margin in the Ethos emulator. They concluded that the reduced margin significantly improved critical OAR dosimetry, which may lead to improved quality of life while preserving or improving tumor coverage [292]. A similar study for anal cancer by Astrom et al. also demonstrated the feasibility of margin reduction with online ART [293]. Other ART strategies have been demonstrated also effective for margin reduction in HNC [294]. These first pieces of evidence indicate that reducing PTV margins can be safely applied in online ART. Future research should address the feasibility of an equivalent approach for IMPT plans optimized with reduced robustness parameters in the online APT context.

Compared to the PTV margin, the possible reduction of the CTV margin during treatment plan adaptation due to tumor shrinkage remains controversial [295]. Even if tumor regression can be imaged, it is not clear how tumor regression is translated at the microscopic level [133]. Hamming et al. showed that adapting to primary tumor GTV shrinkage on MRI mid-treatment was potentially unsafe [296].

Potential for dose escalation

Higher radiotherapy dose intensity has been associated with improved local-regional control and patient outcome [297, 298]. However, dose escalation should be performed while maintaining at least the same dose to the OAR to avoid a higher toxicity [299]. Therefore, dose escalation studies should be carefully addressed. The randomized trial (RTOG-0617) demonstrated that lung cancer patients who received an escalated dose (74 Gy) to the target had worse survival than patients who received the standard dose of 60 Gy [300]. A single institution analysis of 251 NSCLC patients showed that keeping the heart V50 Gy below 25% was associated with a nearly 20% absolute improvement in 2-year overall survival [301]. The more the dose to healthy tissue can be spared, the more room for dose escalation. Integrating proton planning and online APT to ensure adequate daily doses opens the path to safely increase local control rates for patients with high loco-regional failure risk [302, 303].

Enabling soft tissue repositioning

Rigid registration between daily images and the pCT is often performed for patient repositioning based on anatomical structures [304] like bony anatomy (e.g., vertebrae) [79, 305]. This technique is convenient because bone anatomical structures can be easily matched due to their high contrast in CT, CBCT, and orthogonal X-ray images. Moreover, they ensure a correct alignment of the bones with respect to proton beam paths which prevents dose distortions caused by critical density changes. Alternatively, soft-tissue re-positioning has also been suggested for image-guided radiotherapy. Moller et al.[306] demonstrated that the implementation of soft-tissue repositioning combined with adaptive IMRT decreased the dose to the lungs. This gain could be even higher when using IMPT instead of IMRT. However, soft-tissue image guidance in proton therapy is theoretically more challenging due to proton range errors. For instance, Maeda et al. evaluated image-guided proton therapy (IGPT) using an in-room CT. They reported that prostate matching reduced the dose to the rectum and improved target coverage compared to bone matching. However, an adaptive approach was desirable to maintain better dose coverage of the seminal vesicles [128]. Depending on the repositioning strategy, differences in target coverage and adaptation rates were also found for esophageal cancer treated with IMPT [307]. Online (D)APT could facilitate IGPT independently of the registration method (soft-tissue or bony anatomy) used for patient re-positioning. The change of densities in the beam path, such as the entrance of a bone structure, and the daily anatomical information would be considered during plan re-optimization on the image of the day.

6.3 Towards the clinical implementation

Several questions arise when considering the future implementation of online APT. In this section, the ones estimated as more relevant will be addressed.

How much of a priority it is for clinicians?

The POP-ART study interviewed 177 radiation therapy centers across 40 countries to retrieve information about the status of clinical practice regarding ART [160]. Several of the centers interviewed by the POP-ART survey wished to increase their use of ART or change their technique for the two locations addressed in this thesis, head and neck (19%) and lung (13%) cancer. In addition, 14% and 12% of respondents were not applying ART but wished to implement it as a priority for head and neck and lung cancer, respectively. The main barriers hampering the implementation reported in this study were the lack of human and material resources and the technical limitations. One could assume that the demand is expected to be even higher for IMPT due to its increased need for treatment plan adaptation. Therefore, the proposed online APT solutions fully automated on a commercial TPS could bring a promising opportunity to ease the workload of the adaptive workflow and expand the use of APT. The reduction of offline APT and the promises of DAPT could even impact the deployment of proton therapy itself. For instance, clinics with limited resources could be interested in such approaches to expand the locations they are able to treat with protons.

When treatment plan adaptation should be triggered?

Some patient groups and treatment schedules can be expected to benefit more from online adaptation. It is the case of hypo-fractionated treatments, as there is less chance for random daily errors to be averaged out in other treatment fractions. Moreover, a higher daily dose delivered per fraction increases the need for healthy tissue protection. Treatment sites for which density changes frequently occur (e.g., thoracic region) could also benefit more from online APT. Online APT is also an attractive option for pediatrics. Due to their longer survival, children are at a higher risk of secondary cancers caused by doses to the healthy tissues. Online APT can reduce the dose to OAR and thereby reduce the risks. Another possibility would be to use NTCP model-based selection to identify the patients that would benefit the most from online adaptation. Besides establishing whether a patient will benefit from plan adaptation, the need and optimal adaptation frequency should also be determined. A recent study from the Mayo Clinic, analyzed 711 patients treated at their institution with robustly optimized IMPT plans and 2196 verification CTs [26]. Malignancies representing the largest proportion of re-planning were HNC (54%) and lung or chest (58%). However, re-planning rates

vary a lot across studies. Moreover, how the need for adaptation is linked to the planning strategy, clinical experience, and other treatment uncertainties remains under research.

In a recent work by Scandurra et al. [123], only 2 out of 25 nasopharyngeal patients ($\sim 8\%$) required offline plan adaptation when robustly optimized against setup (5mm and 3mm) and range errors (3%). On the contrary, other studies reported much higher adaptation rates for robustly optimized (3mm, 3.5%) IMPT plans in HNC patients. For instance, Gunn et al. [308] found 19 out of 50 (38%) robustly optimized (3mm, 3.5%) oropharyngeal patients to require adaptive re-planning because of weight loss and tumor volume changes. Using the same robustness parameters, Yang et al. reported also that 40% of all head and neck patients treated at their institution needed at least one plan adaptation and 10% more than one [125]. Other small cohort studies have reported inadequate target coverage with no adaptation, 60% by Lalonde et al. [126] and 25% by Cubillos-Mesias [234]. Our findings with 10 oropharyngeal patients robustly optimized (4mm, 2.6%) fit within these intervals. As discussed in Chapter 5, we found that deformed target volumes did not achieve adequate target coverage in 51% of the corrCBCTs. When evaluating the not adapted accumulated doses, the need for adaptation was 30% (3 out of 10 patients). Moreover, when analyzing fraction by fraction, the need for adaptation was reduced at least by half when applying online APT strategies in the fractions where target coverage was not achieved. From these diverse results, it seems unclear and somewhat unpredictable which HNC patients may require plan adaptation and at which point it should occur during their treatment. Therefore, the possibility of performing online APT in a fully automated way could help to adequately deliver IMPT treatments without needing tedious offline APT.

A similar situation is found in lung cancer patients. A 61% was the adaptation rate reported by Hoffman et al. [22] when using a 6mm PTV margin. Szeto et al. [30] found that 50% (8/16) of their lung cancer patient cohort needed adaptation. In a group of 18 patients, Van der Laan et al. [130] evaluated robustly optimized IMPT plans (6mm, 3%) on weekly average 4D-CTs and detected 8 ($\sim 45\%$) patients with re-planning indication, of which 5 patients needed re-planning once and 3 patients more than once. On the contrary, when using the same robustness parameters, Ribeiro et al. [79] reported a lower adaptation rate with only 2 out of 10 (20%) lung cancer patients being re-planned. Similar rates were found in a clinical trial of combined high-dose protons with chemotherapy where 9 out of 44 ($\sim 20\%$) patients benefit from APT. According to our study in 14 lung cancer patients planned with robustly optimized IMPT (5mm, 3%, 3 breathing phases), 35% (5/14) required treatment plan adaptation when evaluating rCTs.

The dose distortions could be compensated with dose restoration to ensure target coverage for all patients except for one presenting non-rigid anatomical changes.

Bobic et al. reported that DAPT could mitigate random delivery errors more effectively than weekly APT on a fraction-per-fraction basis. However, for treatment sites with small day-to-day variations but significant weekly changes, weekly plan adaptation could be sufficient and would reduce the clinical workload [200]. Combinations of different mitigation approaches could also be considered, such as daily dose restoration combined with weekly full (offline) treatment plan adaptation.

Another question that arises is whether DAPT is necessary, even if feasible. Should we perform daily adaptation just because we can? For instance, as reported by Scandurra et al. [123], clinical goals were achieved with robustly optimized IMPT plans and relatively low offline adaptation rates. The current studies with low offline adaptation rates, therefore, question the necessity of DAPT unless it comes with increased dose conformity (e.g., reduction of robustness parameters or CTV-PTV margins). To determine whether DAPT is truly necessary, we need to evaluate if the dosimetric improvements provided by DAPT are clinically relevant for specific planning strategies. Relevant clinical information will arise from the clinical experience with MR-Linacs and the Ethos system, for the moment DAPT has not been clinically implemented. In Chapter 5, we reported improved NTCP values compared to not adaptation for both dysphagia and xerostomia already when applying OAML only to fraction were adequate target coverage was not correct. Higher benefits (lower NTCP and reduced toxicities) could be observed in oropharyngeal patients when combining DAPT and margin reduction or when applying APT based on OAR constraints. Moreover, Nennof et al. [309] compared TCP and NTCP in not adapted plans to DAPT plans for a small lung cancer cohort of 6 patients. Two patients showed substantial TCP reductions up to 33% without adaption. The study concluded that most patients would have benefited from DAPT, either by recovering the TCP or reducing NTCP. With DAPT, one patient showed improvement in predicted 2-year survival based on heart toxicity.

How to trigger online treatment plan adaptation?

Instead of DAPT, an adaptive workflow in which only some fractions would be adapted could be proposed for the first introduction of online APT. Such a workflow requires a strategy to trigger fractions needing adaptation, avoiding false negatives properly. A false negative would mean missing a treatment plan adaptation when actually necessary. On the contrary, If DAPT does not have a significantly increased workload, false positives (adapting when not necessary)

could be tolerated in an ideal DAPT workflow.

In order to trigger treatment plan adaptation, Traffic Light Protocols (TLP) have been developed to recognize and react to anatomical changes. TLP involves Radiation Therapists (RTTs) alerting the radiation oncologists (RadOncs), who should provide the final decision based on the handed-over findings. TLPs have been clinically implemented for offline ART for the locations explored in this thesis: lung [310] and head and neck [117]. However, the TLP workflow can be quite labor-intensive for the RadOncs and the RTTs if the full workflow is triggered even if final action is not required. Alternatively, Buijs et al. [311] proposed the Take Action Protocol (TAP). In this protocol, the RTTs not only had a role in detecting anatomical changes but also decided on the appropriate action and follow-up, resulting in a significant shift in responsibility. These protocols could be combined with dosimetric evaluations on deformed contours as Taasti et al. proposed for offline APT [287] to help decision-making in an online setting.

One of the natural steps forward for this thesis should be to develop a triggering protocol that not only indicates when to adapt but also which online APT technique to use. The design of the protocol would be directly linked to the objective of the treatment plan adaptation. Following the terminology described by Heukelom et al. [146] and detailed in Table 6.1, dose restoration techniques could be used for $ART_{ex-aequo}$ and ART_{OAR} aiming at maintaining target coverage and keeping (or even reduce) the dose to OAR. Alternatively, when seeking to escalate the dose to the target volume or to maintain the dose in a shrunken target volume, full online re-planning relying on deformed contours (e.g., OADEF and OAML) would be necessary for ART_{amplio} , ART_{reduco} and ART_{totale} .

What the role of the radiation oncologist will be?

The presence of the RadOnc in the room at the moment of adaptation is a central point of discussion of the adaptive workflow. In the past, due to the low availability of RadOncs, researchers focused on the complete automation of the adaptive workflow. The claim was that there was no physical time to review and correct the planning contours in an online setting. Following this reasoning, the online APT methods described in this thesis focused on developing fully automated workflows to generate clinically acceptable plans without user alterations. Thus, the proposed solutions aim to avoid or at least minimize the two main actions requiring RadOnc actions: (1) the manual contour delineation at every fraction and (2) the usual physician-physicist loop for plan improvement, review, and approval. Additionally, all the proposed strategies were fully integrated and automated in a commercial TPS.

Name	Technique	Tumor dose	OAR dose
$ART_{ex-aequo}$	Serial plan verification to stabilize initial plan parameters	=	=
ART_{OAR}	Conserve CTV coverage and reduce OAR dose	=	↓
ART_{amplio}	Increased dose to the tumor isotoxic OAR dose	↑	=
ART_{reduco}	"Shrinking CTV" for on-treatment responders	=	↓
ART_{totale}	Increased dose to a shrunken CTV	↑	↓

Table 6.1: Adaptive radiotherapy terminology as proposed by Heukelom et al. [146]

Dose restoration was initially proposed to tackle (1) and (2) simultaneously. The use of not corrected rigidly mapped contours, together with the reconstruction of an already clinically approved dose distribution, aimed at minimizing the involvement of the medical doctor. One could imagine that a trained RTT could verify the restored dose (i.e., the initial dose) on the repeated image. With the arrival of ART clinical tools such as the Ethos system, there has been a paradigm shift in online ART. The emphasis is no anymore on reducing the presence of the RadOnCs in the treatment room but on assisting them at the machine. The correction of automatically generated contours by the RandOnc was the compromise that made possible the integration of online ART in commercial solutions for their clinical usability. This approach, a full new treatment plan based on daily generated (and verified) contours, is the state-of-the-art for online ART as clinically implemented in the Ethos system and MR-Linacs. The disadvantage of this method, and our proposed methods OADEF and OAML, is that they require daily contour verification. AI tools for auto-segmentation have demonstrated higher performance than DIR-based automatic propagation [284]. However, manual checks prior to plan adaptation, and in many cases, minor modifications are still required today [194, 137]. In order to optimize the workflow, a smarter AI contouring protocol was proposed by Chavez-de-Plaza et al. to enable targeted inspection providing AI uncertainty and clinically relevant features (e.g., the proximity of the OAR to the target) [312]. If such methods can be run in a limited time, then applying full online APT could be feasible as it was demonstrated in this thesis.

RadOncs will play a key role in designing the adaptation protocols they want and can clinically implement. This decision will depend on the available tools and human resources. We provided several tools among which physicians, physicists, and RTTs can decide, depending on their clinical objectives and possibilities.

Dose restoration solutions seem to be appropriate tools for clinics aiming to reduce offline adaption specific to protons but lacking resources to implement more complex online APT solutions. For instance, they could apply dose restoration for fractions with density anatomical changes and reserve offline adaptation for critical non-rigid anatomical changes. Another possibility to implement DAPT could be combining dose restoration with full online APT (i.e., OADEF or OAM) for fractions with challenging anatomical changes. Although in Chapter 5 OAML reported the best OAR sparing, this strategy depends on site-specific dose prediction models. The ability to build consistent databases for training and commissioning ML models also needs to be considered. Initially, this might discourage some clinics that prefer generic solutions closer to tools they already master, such as DIR. In those cases, online APT using OADEF could be an interesting alternative.

The role of the RTTs will play a crucial role in the efficient implementation of online APT. An RTT-led staffing model seems the more reasonable pathway to overcome the barrier of RadOncs availability in the adaptive workflow [313, 314]. In the Netherlands, Daal et al. [315] (Ethos system) and Willigenburg et al. [316] (MR-Linac) described the clinical implementation of practices to enable RTT-led online ART in prostate cancer. Physicians can mentor RTTs through traditional tasks previously reserved for the RadOncs, such as contouring (target and OAR volumes) and plan approval achieving comparable outcomes [317]. However, this responsibility shift from the RadOncs towards the RTT should come with a clear definition of competencies. For this purpose, a new figure known as *RTT advanced-adapter* has been proposed for online ART [314]. Comprehensive RTT training programs have already been developed [313, 314], incorporating anatomy contouring and planning training using both emulated environments and real-time on-couch treatments. They indicate that trained RTTs could independently treat prostate patients after the first few fractions. The RadOnc was present for the first week to oversee potential clinical issues and remained *on-call* for the rest of the fractions. A similar strategy could be easily adapted to implement RTT-led online APT. However, legal constraints may differ across countries hampering or easing the clinical implementation depending on how strict legislation is and how fast it could be adapted to the clinical needs. For instance, nowadays in Belgium, a RadOnc and a medical physicist are required to give their approval every time a new adapted plan is generated on the Ethos system.

Are we fast enough?

Reported median times for treatment plan adaptation for the Ethos system in the pelvic region are around 20 minutes [194, 137] with some fractions going up to 40 minutes and even close to one hour for MR-Linacs [191]. Such timing intervals between in-room imaging acquisition and treatment delivery could be even higher for more challenging locations with a larger number of OAR, such as the head and neck location. Moreover, online APT density changes during the treatment plan preparation span could also have a greater impact undermining the benefit of adaptation. The proposed strategies in Chapter 5 provided an adapted plan in around 20 minutes. However, this time accounts only for treatment plan optimization and no-contour verification. This process should be even faster to be applied in a realistic clinical setting. The use of GPUs and the increased computational power that is a reality in the radiation therapy field would very likely ease the timing restrictions. Additionally, the use of analytical algorithms, no correction of deformed structures, or scenario reduction are option to speeding up the process.

Uncertainties in the dose distribution due to anatomical changes have been shown to outweigh differences between analytical and Monte Carlo dose calculation [198]. Moreover, Nenoff et al. also demonstrated in a set of 5 lung cancer patients that DAPT was beneficial even if optimized on uncorrected propagated structures [318]. For an efficient DAPT workflow, another possibility to speed up the plan generation would be to exclude the setup error scenarios from the robust optimization. With a high-quality image of the day (such as in-room CT) and repositioning techniques, this approach could improve speed and OAR sparing due to increased dose conformity. However, Berget et al. demonstrated in a plan-library online APT approach that intrafraction anatomic changes in cervical cancer had a substantial dosimetric effect on target coverage with reduced PTV margins. They suggested that compensation for the uncertainties was still required by using margins or robust optimization for patients with large uterus motions [319].

6.4 Future innovations for online adaptive IMPT

Daily adapted IMPT plans obtained from scratch are expected to become available with further improvements in imaging techniques, dose calculation, and optimization speed. Running the calculations on a GPU and making certain assumptions and simplifications, Matter et al. showed that a simple treatment plan could be generated within 10 seconds [197]. The integration of AI solutions for synthetic CT generation, auto-contouring, and auto-planning on commercial software will play a key role, as discussed in this thesis and demonstrated on the

Ethos system. Nevertheless, for an efficient implementation of online APT, the competencies and organization of the radiation therapy departments will need to be adjusted.

Moreover, DAPT could substantially reduce treatment uncertainties. New daily plans could be adapted to inter-fractional and even intra-fractional anatomical changes to achieve real-time image-guided adaptive treatment delivery. Some remaining uncertainties will still need to be considered in the robust optimization. For instance, the delineation errors, the HU conversion uncertainties (although reduced by dual-energy CT), or some residual setup errors could not be completely removed using APT. However, the uncertainty reduction could enable tighter PTV margins (or reduces robustness parameters), higher dose conformity to the tumor, better healthy tissue sparing, and reduced toxicities. The integration of MRI with proton therapy delivery has been foreseen to be the next major advancement in proton therapy practice [320]. An integrated approach could bring together the excellent real-time soft tissue imaging of MRI with the superior proton dose conformity to accurately compensate inter- and intra-fractional variations. For liver tumors, Moteabbed et al. [321] found that the benefits of MR-integrated proton therapy in liver tumors (e.g., 40% margin reduction and a clinically significant decrease in NTCP) exceed those of currently available image-guided techniques such as MR-linacs, CBCT-guided and offline MR-guided proton therapy.

Although this thesis addressed IMPT delivery, new proton delivery techniques are being explored. Currently, a popular research field for this application is flash therapy which aims at treating tissue with a very high dose rate to induce into the healthy tissue the protective so-called 'Flash-effect' [322, 323]. Current pre-clinical tests are using a ridge filter to create a spread-out Bragg peak for a single beam energy. However, the ridge filter in a flash therapy plan will be much more challenging to adapt online than a multi-layer proton plan, as it needs to be 3D printed. Other novel proton therapy delivery techniques relying on multi-energetic Bragg peaks would still benefit from online APT. It is the case of proton mini-beams [324], proton grid therapy [325], and also proton arc therapy [326].

Regarding treatment fractionation, part of the radiation therapy field is moving towards hypo-fractionated treatments, being the most common treatment sites prostate, lung, and liver [11]. In hypo-fractionated treatments, higher daily doses are delivered to the patient in fewer fractions. The reduced number of fractions can minimize patient burden and treatment costs, while higher doses can be more effective against radio-resistant tumors. The protective effect of proton therapy on normal tissue could increase the toleration for hypo-fractionation. However,

distortions in proton doses induced by anatomical changes are more critical when only a few fractions are delivered. Thus, a higher necessity for APT is expected for very conformal IMPT hypo-fractionated treatments.

Treatment plan adaptation could be also considered in the future not only to compensate anatomical changes but also to account for additional biological information. Functional imaging techniques, including positron emission tomography (PET) combined with CT or MRI, can be used non-invasively to retrieve information about biomarkers such as hypoxia, cell proliferation and apoptosis [327]. These biologic markers can be monitored during the treatment course to guide treatment plan adaptation not only according to anatomical changes but also to the tumor response [328]. This has the potential to advance radiation therapy treatments towards a high level of personalized medicine, with the realistic aim of improving cancer treatment.

In conclusion, this thesis adds to the growing literature indicating that online APT is within reach and that the transition to clinical implementation is feasible in the near future. Ultimately, the broad deployment of online adaptation should meet the objective of reduced radiation-induced side effects and improved quality of life of the treated patients.

Appendix A

Supplementary material Chapter 2

Patient	Location	Nodal involvement	CTV tumor size [cm ³]			Tumor motion amplitude [mm]		
			pCT	$\Delta 1rCT$	$\Delta 2rCT$	R-L	A-P	S-I
1	RML	N	152.6	15.8	-16.1	4.2	2.1	3.1
2	LUL	Y	165.1	-24.2	-26.5	1.1	1.1	2.0
3	RML	Y	331.2	-112.4	-133.1	0.4	0.5	1.9
4	RUL	Y	146.4	-86.1	-0.7	1.2	0.7	1.6
5	RUL	Y	343.4	-3.9	-58.6	0.4	1.4	0.5
6	LLL	Y	177.6	-52.8	-61.0	3.5	3.0	3.4
7	RUL	Y	79.1	12.4	4.0	1.3	3.5	0.7
8	RUL	Y	133.7	-39.0	-32.2	0.9	0.5	2.9
9*	LUL	Y	442.5	-141.8	No data	0.7	0.8	0.3
10	LUL	Y	130.2	-49.4	-47.6	0.8	0.9	1.5
11	RUL	N	109.6	-12.5	-29.6	2.2	1.8	6.6
12	RLL	Y	294.4	-5.1	-32.2	2.1	2.5	10.6
13	RML	Y	218.1	-29.0	-44.6	1.0	3.2	4.9
14	RUL	Y	268.1	-70.5	-114.9	4.3	2.8	3.0

Table A.1: Patient characteristics. Size evolution on repeated images ($\Delta 1rCT$, $\Delta 2rCT$) is given respect to the initial volume. Tumor motion amplitude is given in absolute value. Abbreviations: LLL = left lower lobe; LUL = left upper lobe; RLL = right lower lobe; RML = right middle lobe; RUL = right upper lobe; Y = Yes; N= No; pCT = Planning CT; rCT = Repeated CT; R-L = Right-Left; AP = Anterior-Posterior; SI = Superior-Inferior; *Only one repeated 4D-CT (2rCT is missing); X=CTV_{map}; Y=CTV_{real}.

Patient	Dice Similarity Coefficient (DSC) $2 X \cap Y /(X + Y)$	
	1rCT	2rCT
1	0.86	0.85
2	0.74	0.80
3	0.75	0.71
4	0.82	0.73
5	0.83	0.88
6	0.36	0.32
7	0.81	0.82
8	0.70	0.67
9(*)	0.82	No data
10	0.71	0.67
11	0.84	0.80
12	0.78	0.79
13	0.76	0.77
14	0.70	0.68

Table A.2: Dice coefficients between mapped CTVs and real CTVs where calculated for all patients in both repeated-CTs. Dice coefficient equal to 1 correspond to perfect overlapping between structures. which is impossible because of tumor growing/shrinkage. However. the closer the Dice coefficient to zero. the less structures are overlapping. meaning a lower performance of the registration strategy. Abbreviations: (*)Only one repeated 4D-CT (2rCT is missing); $X=CTV_{map}$; $Y=CTV_{real}$; 1rCT = first repeated CT; 2rCT = second repeated CT.

Appendix B

Supplementary material Chapter 3

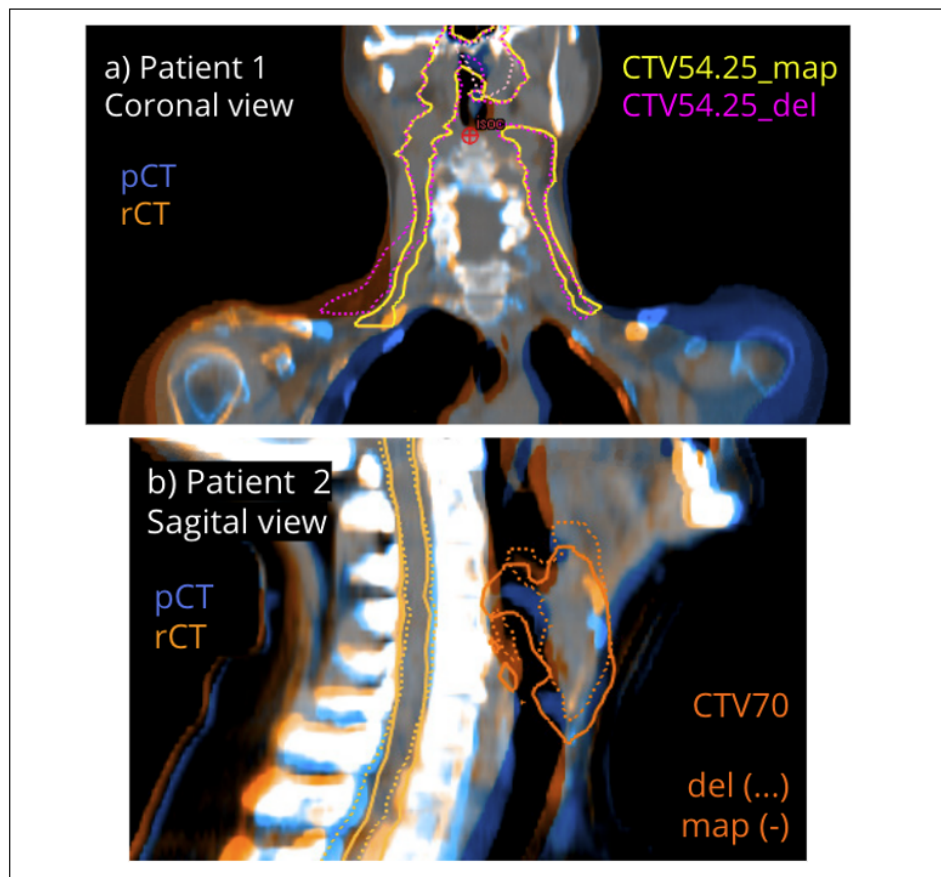


Figure B.1: Views of rigid registrations between planning-CT (pCT) and 5th-repeated-CT for patient 1 (a) and between planning-CT and 3rd-repeated-CT for patient 2 (b). Rigidly mapped contours from pCT to repeated-CTs (rCTs) are represented with solid lines while delineated contours on rCTs are shown with dotted lines. pCTs and rCTs are blue and orange colored, respectively.

NOMINAL scenario			
ROI(Metric)	NA/DR	NA/FOA	DR/FOA
CTV54.25 (D98%)	2,12E-04	2,51E-11	3,70E-13
CTV70 (D2%)	1,32E-04	8,29E-06	1,82E-02
CT70 (D98%)	2,76E-07	7,60E-13	4,19E-02
Parotid_R(DMean)	7,28E-01	1,45E-01	1,40E-01
Parotid_L(DMean)	4,22E-01	6,88E-03	3,46E-02
Right Submandibular (DMean)	6,61E-01	9,62E-01	6,90E-01
Left Submandibular (DMean)	3,16E-02	3,72E-01	8,14E-01
SpinalCord (D0.03cm3)	5,27E-04	4,14E-04	6,87E-02
Oral Cavity(DMean)	4,60E-01	8,12E-01	3,61E-01
WORST-CASE scenario			
ROI(Metric)	NA/DR	NA/FOA	DR/FOA
CTV54.25 (D98%)	7,80E-04	8,86E-05	1,20E-04
CTV70 (D2%)	3,31E-03	1,32E-04	7,66E-02
CT70 (D98%)	3,78E-04	4,01E-05	2,00E-01
Right Parotid (DMean)	4,90E-02	8,00E-01	8,64E-02
Left Parotid (DMean)	5,89E-02	4,71E-01	4,51E-01
Right Submandibular (DMean)	3,23E-01	4,64E-01	8,27E-01
Left Submandibular (DMean)	1,32E-01	4,94E-01	6,15E-01
Spinal Cord(D0.03cm3)	1,10E-06	1,03E-04	7,43E-01
Oral Cavity(DMean)	1,31E-01	2,86E-01	1,46E-01

Table B.1: P-values for each clinical goal for the significant Wilcoxon ranged paired test between the three compared plans: not-adapted robustly optimized clinical plans (NA), restored plans (DR) and fully offline adapted plans (FOA). Statistically significant values ($p < 0.05$) are colored in green. Metrics with no statistically significant values for any of the plan combinations are highlighted in grey.

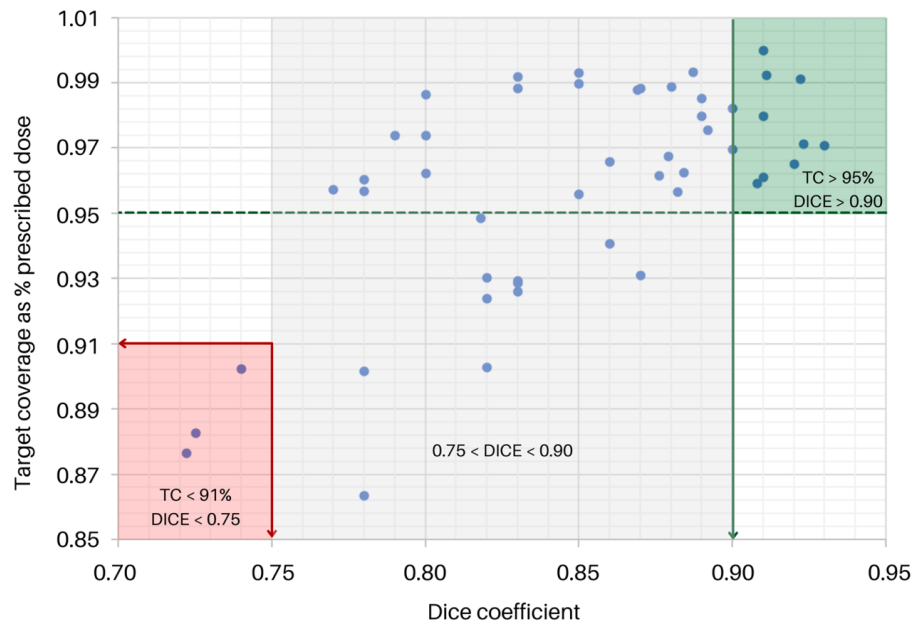


Figure B.2: Target coverage of the CTV in % of its prescribed dose plotted against Dice Similarity Coefficient (DSC) between rigidly mapped and manually delineated CTV. Optimal (green) and sub-optimal (red) target coverage correspond to high DSC coefficient (>0.90) and low DSC coefficients (<0.75) respectively.

Appendix C

Supplementary material Chapter 4

SM1. Proton therapy Pencil Beam Scanning (PBS) for HN cancer (PROKET-MIROv8)

Sixty local oropharynx cancer patients were consistently planned by the same radiation therapy technician (RTT). Average high-risk clinical target volume (CTV) and low-risk CTV volumes were 133 cc (range: 10–300 cc) and 392 cc (range: 274–708 cc), respectively. The prescribed doses were 70Gy to the primary tumor (high-risk CTV) and pathological lymph nodes, and 54.25Gy to the elective nodal areas (low-risk CTV), to be delivered in 35 fractions. The plans were generated through multiple robust optimization iterations (range error = 2.6%, setup error = 4 mm) with 21 (3×7) uncertainty scenarios and a final Monte Carlo (MC) dose calculation with an uncertainty of 1%. The dose grid was 3×3×3 mm³. Below we detail the steps followed to generate the plans:

- First optimization round focusing on the CTV coverage with 150 iterations and 5000 ions per spot
- Robustness evaluation run on 2 (range error) × 14 (setup error) = 28 scenarios, accounting for ±2.6% proton range error and ±4mm systematic error in the 3 directions of the space (6 points) and the diagonals (8 points).
- Then, OAR clinical objectives were included: First focus on clinical goals with priority 1 and 2, body Dmax for the hotspots and robustness in target coverage.
- Re-evaluation of the robustness of the plan and fine tuning to avoid hotspots
- Refining the iterations by considering all OARs, especially the ones with Dmean constraints
- For the final plan, final robust evaluation including a third discretization point for the range uncertainty (3×14 = 42 scenarios)

The list of clinical goals to accomplish with their priority is detailed in Table C.1

Priority	Region of interest (ROI)	Clinical goal
1	Low-risk CTV	D98% >51.54 Gy (95%Dp = 54.25 Gy)
1	High-risk CTV	D98% >66.50 Gy (95% Dp = 70 Gy)
1	High-risk CTV	D0.03 cm3 >74.90 Gy (107% Dp = 70 Gy)
1	Body	D0.03 cm3 >74.90 Gy (107% Dp = 70 Gy)
1	Spinal Cord	D0.03 cm3 <45 Gy
1	Brain Stem	D0.03 cm3 <54 Gy
2	Low-risk CTV	D98% >53.17 Gy (98% Dp = 54.25 Gy)
2	High-risk CTV	D98% >68.60 Gy (98% Dp = 70 Gy)
3	PCMSup/PCMMid/PCMInf	Mean dose <40 Gy
3	Parotids (left and right)	Mean dose <26 Gy
3	SMGs (left and right)	Mean dose <40 Gy
3	Oral cavity	Mean dose <35 Gy
3	Larynx	Mean dose <50 Gy
3	Mandible	D0.03cm3 <70 Gy
3	Glottic area	Mean dose <40 Gy
3	Esophagus	Mean dose <35 Gy
3	Choclea (left and right)	Mean dose <30 Gy

Table C.1: List of clinical goals from the planning protocol used in this study to generate the ground truth treatment plans. Abbreviations: PCM = pharyngeal constrictor muscles, Sup = Superior, Mid = Mid = Middle, Inf = Inferior, SMG: submandibular glands

**SM2. Post-processing parameters for the mimicking model
in RSL-IMPT-Oropharynx-7000-SIB (3.0)**

Functions for target volumes			
Roi	GoalType	Function	Value (Gy)
CTV70	MinDose	Aggressive	69.50
CTV54.25	MinDose	Aggressive	53.75
CTV54.25+13mm	MaxDose	Aggressive	55.00
-CTV70+15mm	MaxDose	Aggressive	71.00
CTV70+6mm	MaxDose	Aggressive	69.50
CTV70	MaxDose	WeightedDvhShift	(vol threshold 0.5)
CTV54.25+3mm	MaxDose	Aggressive	69.00
-CTV70	MaxDose	Aggressive	68.00
CTV54.25+3mm	MaxDose	Aggressive	66.50
-CTV70+3mm	MaxDose	Aggressive	54.25
CTV70+13mm	MaxDose	Aggressive	54.25
-CTV70+3mm	MaxDose	Aggressive	53.50
CTV54.25+13mm	MaxDose	Aggressive	51.50
-CTV54.25+8mm&CTV70+18mm	MaxDose	Aggressive	
CTV54.25+5mm-CTV70+4mm	MaxDose	Aggressive	
CTV54.25+5mm-CTV70+20mm	MaxDose	Aggressive	
CTV54.25+13mm	MaxDose	Aggressive	
-CTV54.25&CTV70+15mm	MaxDose	Aggressive	
Functions for OAR volumes			
Roi	GoalType	Function	Value
Parotid L/Parotid R	Reduce Dose	-	0.15
Oral Cavity	Reduce Dose	-	0.2
SMG L/SMG R	Reduce Dose	-	0.1
PCM Inf/PCM Mid/ PCM Sup	Reduce Dose	-	0.2
Esophagus	Reduce Dose	-	0.2
PCM(Inf/Mid/sup)	MaxDose	Aggressive	62.00
-CTV70+6mm	MaxDose	Aggressive	54.25
PCM(Inf/Mid/sup)	MaxDose	Aggressive	54.25
-CTV70+6mm	MaxDose	Aggressive	54.25
OralCavity-CTV70+6mm	MaxDose	Aggressive	54.25
Parotid (R/L)-CTV70+6mm	MaxDose	Aggressive	54.25
SMG (R/L)-CTV70+6mm	MaxDose	Aggressive	54.25
Spinal Cord	MaxDose	Aggressive	40.00
Spinal Cord+3mm	MaxDose	Aggressive	41.00
Brain	MaxDose	Aggressive	60.00
BrainStem	MaxDose	Aggressive	60.00

Table C.2: List of model parameters for the post processing applied to the predicted dose distribution. More information about the parameters can be found in the guide “RSL-D-RS-11B-SEML-EN-1.0-2021-12-10 RAYSTATION 11B SCRIPTING ENVIRONMENTS FOR MACHINE LEARNING” pages 12,13.

SM3. Robust optimization for the dose mimicking methods

The single predicted dose (with or without post-processing) is used to guide the standard robust optimization as usually performed in RayStation (worst-case optimization). The basis of this method is minimax optimization, in which the optimization functions that have been selected to be robust are considered under the worst-case scenario C.1. The worst-case scenario is the realization of uncertainty under which a robust function attains its highest value. If several functions are selected to be robust, their weighted sum in the worst-case scenario is considered. Minimax optimization of n functions $f_1, \dots, f_n$ which are all required to be robust over the scenarios enumerated by the set S , and which have non-negative importance weights $w_1, \dots, w_n$, can be formulated as an optimization problem on the form:

$$\min_{x \in X} \max_{s \in S} \sum_{i=1}^n w_i f_i(d(x; s)) \quad (\text{C.1})$$

where X is the set of feasible variables (spot weights for IMPT), and $d(x; s)$ is the dose distribution as a function of the variables x and the scenario s .

In the case of IBM, the predicted dose information (dose shape) is encoded in the isodose volumes generated from it. For this method, robust optimization is introduced by setting as robust functions the subset of optimization objectives described in Table C.3.

For RSM, the voxel-wise-based optimization objectives provide the contribution of the predicted dose to guide the optimization. During the optimization, voxels that contain dose values very different from the values of the predicted dose are penalized. For RSM, robust optimization is introduced by setting as robust functions the subset of optimization objectives described in Table C.4.

SM4. Optimization objectives used in Isodose Based Mimicking (IBM)

Roi	Objective	Dose value (RBE)	Robust	Weight
CTV high-dose	Min	70.00 Gy	*	700
CTV high-dose	Max	70.00 Gy	*	500
CTV low-dose (right)	Min	54.25 Gy	*	700
CTV low-dose (left)	Min	54.25 Gy	*	700
CTV high-dose + 3 cm	Max	70.00 Gy	*	9000
isodose 7150	Max	72.20 Gy	*	400
isodose 6950	Min	69.50 Gy	*	100
isodose 6950	Max	71.50 Gy		200
isodose 6750	Min	67.50 Gy	*	100
isodose 6750	Max	69.50 Gy		
isodose 6650	Min	66.50 Gy	*	100
isodose 6650	Max	67.50 Gy		10
isodose 6150	Min	61.50 Gy		1
isodose 6150	Max	66.50 Gy		10
isodose 5696	Min	56.96 Gy		1
isodose 5696	Max	61.50 Gy		10
isodose 5496	Min	56.96 Gy		1
isodose 5496	Max	61.50 Gy		10
isodose 5296	Min	52.96 Gy	*	100
isodose 5296	Max	54.96 Gy		50
isodose 5153	Min	51.53 Gy	*	100
isodose 5153	Max	52.96 Gy		20
isodose 4653	Max	51.53 Gy		50
isodose 4153	Max	46.53 Gy		50
isodose 3653	Max	41.53 Gy		50
isodose 3153	Max	36.53 Gy		50
isodose 2653	Max	31.53 Gy		5
isodose 2153	Max	26.53 Gy		10
isodose 1653	Max	21.53 Gy		5
isodose 1153	Max	16.53 Gy		10
isodose 653	Max	11.53 Gy		1
isodose 153	Max	6.53 Gy		5
isodose 0	Max	1.50 Gy		1

Table C.3: List of dose objectives used for worst-case optimization with the isodose-based mimicking (IBM) method. The maximum dose objectives were set as robust in order to avoid hot spots, while the minimum dose objectives aimed at ensuring robustness in the target coverage. The use of dose objectives on the isodose volumes during the isodose-based mimicking should enable the replication of the dose gradient from the predicted dose into the mimicked dose. The star (*) indicates the optimization objectives set as robust.

**SM5. Optimization objectives used in the RayStation mimicking (RSM) model
hardcoded in the model RSL-IMPT-Oropharynx-7000-SIB (3.0)**

Strategies: MimicSettings: Functions				
Roi	Function	Weight	IsoWeights	IsoDose
CTV54.25	MinRefDose	1e-06	None	None
CTV54.25	MaxRefDose	1e-06	1, 1, 1, 0.1	104, 100, 77.5, 0
CTV70+3mm	MinRefDose	1e-06	None	None
CTV70+3mm	MaxRefDose	1e-06	10, 6, 1, 0.1	104, 100, 98, 0
CTV54.25+3mm	MinRefDose	1e-06	None	None
CTV54.25+3mm	MaxRefDose	1e-06	10, 6, 1, 0.1, 10, 0.1, 0.1	104, 100, 93, 85.5, 77.5, 75, 0
Brain	MaxRefDose	0.001	None	None
Brainstem	MaxRefDose	0.3	None	None
SpinalCord	MaxRefDose	0.5	None	None
Parotids/CTVs+3mm	MaxRefDose	0.175	None	None
Parotids-CTVs+3mm	MaxRefDose	0.35	None	None
Parotids-CTVs+5mm	MaxRefDose	0.7	None	None
SMG/CTVs+3mm	MaxRefDose	0.25	None	None
SMG-CTVs+3mm	MaxRefDose	0.5	None	None
SMG-CTVs+5mm	MaxRefDose	1	None	None
OralCavity/CTVs+3mm	MaxRefDose	0.1	None	None
OralCavity-CTVs+3mm	MaxRefDose	0.4	None	None
OralCavity-CTVs+5mm	MaxRefDose	0.15	None	None
Strategies: MimicSettings: RayStationObjectiveFunctions: Function				
Roi	Function	Weight	DoseLevel	RobustRoi
CTV70	MinDose	2000	70.00	*
CTV70+10mm	MaxDose	1000	71.00	*
CTV70+10mm	MaxDose	1000	70.50	
CTV70	UniformDose	750	70.00	
CTV54.25	MinDose	2000	54.25	*
CTV54.25-CTV70+10mm	MaxDose	1000	57.00	*
CTV54.25-CTV70+10mm	MaxDose	1000	57.00	
CTV54.25-CTV70+10mm	MaxDose	75	54.25	*
CTV54.25-CTV70+10mm	MaxDose	75	54.25	
CTV70	MinDose	4000	70.00	*
CTV70+10mm	MaxDose	200000	70.50	*
CTV70+10mm	MaxDose	100000	70.50	
CTV54.25	MinDose	6000	54.25	*

Table C.4: List of model parameters for the RayStation dose mimicking associated to the machine learning model 'RSL-IMPT-Oropharynx-7000-SIB (3.0)' More information about the parameters can be found in the guide "RSL-D-RS-11B-SEML-EN-1.0-2021-12-10 RAYSTATION 11B SCRIPTING ENVIRONMENTS FOR MACHINE LEARNING" pages 14-17. The star (*) indicates the optimization objectives set as robust. Abbreviations: CTV = Clinical Target Volume, SMG = Submandibular Glands

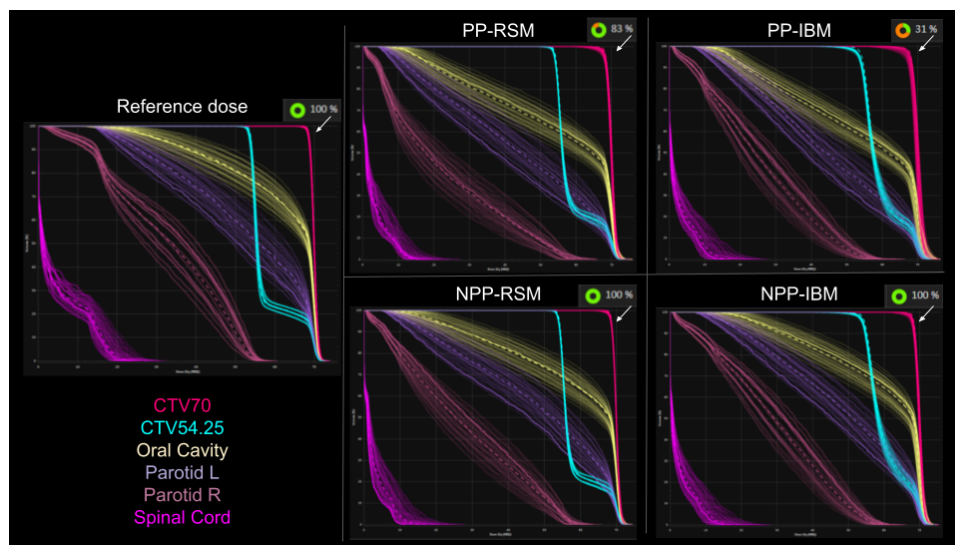


Figure C.1: The differences in terms of clinical planning criteria for the OARs between KBP plans and the clinical plans. Negative differences imply that the KBP plan was better than the clinical plan for that metric (i.e., less dose to the OARs). The boxes indicate the median and interquartile range (IQR). The whiskers extend to a minimum of 1.5 times the IQR and diamonds represent the outliers. Abbreviations: KBP = Knowledge-Based Planning, PP = Post-processing, NPP = No Post-Processing, IBM = Isodose-Based Mimicking, RSM = RayStation Mimicking, SMG = Submandublar Gland, PCM = Pharyngeal Constrictor Muscle

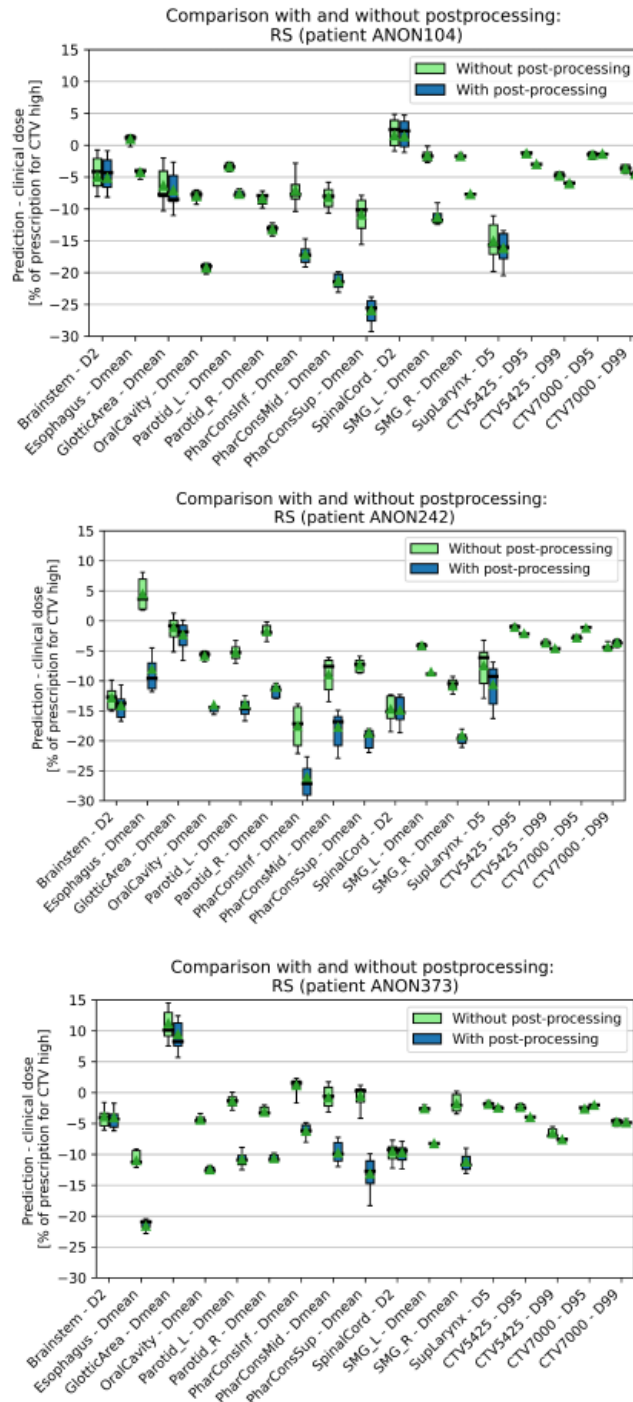


Figure C.2: For the three left out patients the DVH metrics from the 11 dose prediction models are compared to the values of the clinical dose. Each boxplot contains the data points for the 11 models. The bigger the boxplot the higher the variability between models.

References

- [1] Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.
- [2] Is the world making progress against cancer? <https://ourworldindata.org/progress-against-cancer>. Accessed: 2022-10-03.
- [3] S C Hiom. Diagnosing cancer earlier: reviewing the evidence for improving cancer survival. *Br. J. Cancer*, 112 Suppl 1:S1–5, March 2015.
- [4] T P Hanna, J Shafiq, G P Delaney, S K Vinod, S R Thompson, and M B Barton. The population benefit of evidence-based radiotherapy: 5-year local control and overall survival benefits. *Radiother. Oncol.*, 126(2):191–197, February 2018.
- [5] Rachel Abudu, Gauthier Bouche, Karima Bourougaa, Lynne Davies, Kalina Duncan, Carla Estaquio, Anna Diaz Font, Marc S Hurlbert, Paul Jackson, Linda Kroeskop-Bossenbroek, Ian Lewis, Giota Mitrou, Abdul Mutabbir, Christopher A Pettigrew, Lynn Turner, Annemarie Weerman, and Kari Wojtanik. Trends in international cancer research investment 2006-2018. *JCO Glob Oncol*, 7:602–610, April 2021.
- [6] Anna Schmutz, Claire Salignat, Daria Plotkina, Amandine Devouassoux, Teresa Lee, Melina Arnold, Morten Ervik, and Olaf Kelm. Mapping the global cancer research funding landscape. *JNCI Cancer Spectr*, 3(4):kz069, December 2019.
- [7] Josep M Borrás, Yolande Lievens, Peter Dunscombe, Mary Coffey, Julian Malicki, Julieta Corral, Chiara Gasparotto, Noemie Defourny, Michael Barton, Rob Verhoeven, Liesbeth van Eycken, Maja Primic-Zakelj, Maciej Trojanowski, Primož Strojan, and Cai Grau. The optimal utilization proportion

- of external beam radiotherapy in european countries: An ESTRO-HERO analysis. *Radiother. Oncol.*, 116(1):38–44, July 2015.
- [8] Josep M Borrás, Yolande Lievens, Michael Barton, Julieta Corral, Jacques Ferlay, Freddie Bray, and Cai Grau. How many new cancer patients in europe will require radiotherapy by 2025? an ESTRO-HERO analysis. *Radiother. Oncol.*, 119(1):5–11, April 2016.
- [9] Josep M Borrás, Cai Grau, Julieta Corral, Karen Wong, Michael B Barton, Jacques Ferlay, Freddie Bray, and Yolande Lievens. Estimating the number of fractions by tumour site for european countries in 2012 and 2025: An ESTRO-HERO analysis. *Radiother. Oncol.*, 126(2):198–204, February 2018.
- [10] Roger G Dale and Bleddyn Jones. *Radiobiological Modelling in Radiation Oncology*. British Inst of Radiology, 2007.
- [11] Alexandre Santos, Scott Penfold, Peter Gorayski, and Hien Le. The role of hypofractionation in proton therapy. *Cancers*, 14(9):2271, 2022.
- [12] Cynthia L Kryder. Proton-beam therapy versus photon-beam therapy: The debate continues. *Radiation Oncology*, 2017.
- [13] Johannes A Langendijk, Patricia Doornaert, Irma M Verdonck-de Leeuw, Charles R Leemans, Neil K Aaronson, and Ben J Slotman. Impact of late treatment-related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. *Journal of clinical oncology*, 26(22):3770–3776, 2008.
- [14] S Teoh, F Fiorini, B George, KA Vallis, and F Van den Heuvel. Proton vs photon: A model-based approach to patient selection for reduction of cardiac toxicity in locally advanced lung cancer. *Radiotherapy and Oncology*, 152:151–162, 2020.
- [15] A J Lomax. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties. *Phys. Med. Biol.*, 53(4):1027–1042, February 2008.
- [16] A J Lomax. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions, 2008.
- [17] Birgit Sabine Müller, Marciana Nona Duma, Severin Kampfer, Simeon Nill,

- Uwe Oelfke, Hans Geinitz, and Jan Jakob Wilkens. Impact of interfractional changes in head and neck cancer patients on the delivered dose in intensity modulated radiotherapy with protons and photons. *Phys. Med.*, 31(3):266–272, May 2015.
- [18] Harald Paganetti. Range uncertainties in proton therapy and the role of monte carlo simulations. *Physics in Medicine & Biology*, 57(11):R99, 2012.
- [19] Peter Park. Development of a beam-specific planning target volume and a robust plan analysis tool for proton therapy. 2012.
- [20] Jan-Jakob Sonke, Marianne Aznar, and Coen Rasch. Adaptive radiotherapy for anatomical changes, 2019.
- [21] Francesca Albertini, Alessandra Bolsi, Antony J Lomax, Hans Peter Rutz, Beate Timmerman, and Gudrun Goitein. Sensitivity of intensity modulated proton therapy plans to changes in patient weight. *Radiotherapy and Oncology*, 86(2):187–194, 2008.
- [22] Lone Hoffmann, Markus Alber, Maria Fuglsang Jensen, Marianne Ingerslev Holt, and Ditte Sloth Møller. Adaptation is mandatory for intensity modulated proton therapy of advanced lung cancer to ensure target coverage. *Radiother. Oncol.*, 122(3):400–405, March 2017.
- [23] R Melissa Anakotta, Hans P van der Laan, Sabine Visser, Cassia O Ribeiro, Margriet Dieters, Johannes A Langendijk, Stefan Both, Erik W Korevaar, Nanna M Sijtsema, Antje Knopf, et al. Weekly robustness evaluation of intensity-modulated proton therapy for oesophageal cancer. *Radiotherapy and Oncology*, 151:66–72, 2020.
- [24] Joe Y Chang, Heng Li, X Ronald Zhu, Zhongxing Liao, Lina Zhao, Amy Liu, Yupeng Li, Narayan Sahoo, Falk Poenisch, Daniel R Gomez, et al. Clinical implementation of intensity modulated proton therapy for thoracic malignancies. *International Journal of Radiation Oncology* Biology* Physics*, 90(4):809–818, 2014.
- [25] Jaden D Evans, Riley H Harper, Molly Petersen, William S Harmsen, Aman Anand, Ashley Hunzeker, Noelle C Deiter, Heather Schultz, Krishan R Jethwa, Scott C Lester, et al. The importance of verification ct-qa scans in patients treated with impt for head and neck cancers. *International journal of particle therapy*, 7(1):41–53, 2020.

- [26] Yue-Houng Hu, Riley H Harper, Noelle C Deiter, Jaden D Evans, Anita Mahajan, Jon J Kruse, and Daniel W Mundy. Analysis of the rate of re-planning in spot-scanning proton therapy. *International Journal of Particle Therapy*, 9(2):49–58, 2022.
- [27] Haihua Yang, Wei Hu, Wei Wang, Peifang Chen, Weijun Ding, and Wei Luo. Replanning during intensity modulated radiation therapy improved quality of life in patients with nasopharyngeal carcinoma. *International Journal of Radiation Oncology* Biology* Physics*, 85(1):e47–e54, 2013.
- [28] J Castelli, A Simon, C Lafond, N Perichon, B Rigaud, E Chajon, B De Bari, M Ozsahin, J Bourhis, and R De Crevoisier. Adaptive radiotherapy for head and neck cancer. *Acta Oncologica*, 57(10):1284–1292, 2018.
- [29] Ditte Sloth Møller, Christina Maria Lutz, Azza Ahmed Khalil, Markus Alber, Marianne Ingerslev Holt, Maria Kandi, Hjørdis Hjalting Schmidt, Marie Tvillum, Ane Appelt, Marianne Marquard Knap, et al. Survival benefits for non-small cell lung cancer patients treated with adaptive radiotherapy. *Radiotherapy and Oncology*, 168:234–240, 2022.
- [30] Yenny Z Szeto, Marnix G Witte, Simon R van Kranen, Jan-Jakob Sonke, José Belderbos, and Marcel van Herk. Effects of anatomical changes on pencil beam scanning proton plans in locally advanced nscl patients. *Radiotherapy and Oncology*, 120(2):286–292, 2016.
- [31] Tobias Finazzi, Cornelis JA Haasbeek, Femke OB Spoelstra, Miguel A Palacios, Marjan A Admiraal, Anna ME Bruynzeel, Berend J Slotman, Frank J Lagerwaard, and Suresh Senan. Clinical outcomes of stereotactic mr-guided adaptive radiation therapy for high-risk lung tumors. *International Journal of Radiation Oncology* Biology* Physics*, 107(2):270–278, 2020.
- [32] Mojtaba Moazzezi, Brent Rose, Kelly Kisling, Kevin L Moore, and Xenia Ray. Prospects for daily online adaptive radiotherapy via ethos for prostate cancer patients without nodal involvement using unedited cbct auto-segmentation. *Journal of applied clinical medical physics*, 22(10):82–93, 2021.
- [33] Kinga Bernatowicz, Xavier Geets, Ana Barragan, Guillaume Janssens, Kevin Souris, and Edmond Sterpin. Feasibility of online IMPT adaptation using fast, automatic and robust dose restoration, 2018.
- [34] Thyrza Jagt, Sebastiaan Breedveld, Steven van de Water, Ben Heijmen, and

- Mischa Hoogeman. Near real-time automated dose restoration in IMPT to compensate for daily tissue density variations in prostate cancer. *Phys. Med. Biol.*, 62(11):4254–4272, June 2017.
- [35] Chris McIntosh, Mattea Welch, Andrea McNiven, David A Jaffray, and Thomas G Purdie. Fully automated treatment planning for head and neck radiotherapy using a voxel-based dose prediction and dose mimicking method. *Phys. Med. Biol.*, 62(15):5926–5944, July 2017.
- [36] Aaron Babier, Rafid Mahmood, Binghao Zhang, Victor GL Alves, Ana Maria Barragán-Montero, Joel Beaudry, Carlos E Cardenas, Yankui Chang, Zijie Chen, Jaehee Chun, et al. Openkbp-opt: An international and reproducible evaluation of 76 knowledge-based planning pipelines. *arXiv preprint arXiv:2202.08303*, 2022.
- [37] Directory of radiotherapy centers (dirac), status of radiation therapy equipment. <https://dirac.iaea.org/Query/Map2?mapId=2>. Accessed: 2022-11-04.
- [38] Daniela Schulz-Ertner and Hirohiko Tsujii. Particle radiation therapy using proton and heavier ion beams. *Journal of clinical oncology*, 25(8):953–964, 2007.
- [39] Makbule Tambas, Hans Paul van der Laan, Roel JHM Steenbakkers, Jerome Doyen, Beate Timmermann, Ester Orlandi, Morten Hoyer, Karin Haustermans, Petra Georg, Neil G Burnet, et al. Current practice in proton therapy delivery in adult cancer patients across europe. *Radiotherapy and Oncology*, 167:7–13, 2022.
- [40] Pierre Blanchard, Adam S Garden, G Brandon Gunn, David I Rosenthal, William H Morrison, Mike Hernandez, Joseph Crutison, Jack J Lee, Rong Ye, C David Fuller, et al. Intensity-modulated proton beam therapy (impt) versus intensity-modulated photon therapy (imrt) for patients with oropharynx cancer—a case matched analysis. *Radiotherapy and Oncology*, 120(1):48–55, 2016.
- [41] Moataz Reda, Alexander F Bagley, Husam Y Zaidan, and Wassana Yantasee. Augmenting the therapeutic window of radiotherapy: A perspective on molecularly targeted therapies and nanomaterials. *Radiotherapy and Oncology*, 150:225–235, 2020.
- [42] Su I Chung, DeeDee K Smart, Eun Joo Chung, and Deborah E Citrin. Ra-

- dioprotection as a method to enhance the therapeutic ratio of radiotherapy. *Increasing the therapeutic ratio of radiotherapy*, pages 79–102, 2017.
- [43] Jonathan R Hughes and Jason L Parsons. Flash radiotherapy: current knowledge and future insights using proton-beam therapy. *International journal of molecular sciences*, 21(18):6492, 2020.
- [44] Bleddyn Jones and David AL Morgan. 4 radiotherapy fractionation. In *Radiobiological modelling in radiation oncology*, pages 51–78. The British Institute of Radiology, 2007.
- [45] Jacob S Parzen, William Hartsell, John Chang, Smith Apisarnthanarax, Jason Molitoris, Michael Durci, Henry Tsai, James Urbanic, Jonathan Ashman, Carlos Vargas, et al. Hypofractionated proton beam radiotherapy in patients with unresectable liver tumors: multi-institutional prospective results from the proton collaborative group. *Radiation Oncology*, 15(1):1–8, 2020.
- [46] Robert Huddart, Shaista Hafeez, Rebecca Lewis, Helen McNair, Isabelle Syndikus, Ann Henry, John Staffurth, Monisha Dewan, Catalina Vassallo-Bonner, Syed Ali Moinuddin, et al. Clinical outcomes of a randomized trial of adaptive plan-of-the-day treatment in patients receiving ultra-hypofractionated weekly radiation therapy for bladder cancer. *International Journal of Radiation Oncology* Biology* Physics*, 110(2):412–424, 2021.
- [47] Particle therapy co-operative group (ptcog). <https://www.ptcog.ch/index.php>. Accessed: 2022-09-08.
- [48] Teruhiko Terasawa, Tomas Dvorak, Stanley Ip, Gowri Raman, Joseph Lau, and Thomas A Trikalinos. Systematic review: charged-particle radiation therapy for cancer. *Annals of internal medicine*, 151(8):556–565, 2009.
- [49] Piet Dirix and Sandra Nuyts. Evidence-based organ-sparing radiotherapy in head and neck cancer. *The lancet oncology*, 11(1):85–91, 2010.
- [50] P Ahn, A Lin, O Zhou, J N Lukens, and S Sharma. A comparative study of Patient-Reported quality of life, xerostomia, and dysgeusia in oropharyngeal squamous cell carcinoma (OPSCC) treated with volumetric modulated arc therapy (VMAT) or proton pencil beam scanning (PBS), 2016.
- [51] Kristin A Higgins, Kelli O’Connell, Yuan Liu, Theresa W Gillespie, Mark W McDonald, Rath N Pillai, Kirtesh R Patel, Pretesh R Patel, Clifford G

- Robinson, Charles B Simone II, et al. National cancer database analysis of proton versus photon radiation therapy in non-small cell lung cancer. *International Journal of Radiation Oncology* Biology* Physics*, 97(1):128–137, 2017.
- [52] Antony J Lomax, Thomas Bortfeld, Gudrun Goitein, Juergen Debus, Christine Dykstra, Pierre-Alain Tercier, Philippe A Coucke, and Rene O Mirimanoff. A treatment planning inter-comparison of proton and intensity modulated photon radiotherapy. *Radiotherapy and Oncology*, 51(3):257–271, 1999.
- [53] Christine S Chung, Torunn I Yock, Kerrie Nelson, Yang Xu, Nancy L Keating, and Nancy J Tarbell. Incidence of second malignancies among patients treated with proton versus photon radiation. *International Journal of Radiation Oncology* Biology* Physics*, 87(1):46–52, 2013.
- [54] Harald Paganetti. Assessment of the risk for developing a second malignancy from scattered and secondary radiation in radiation therapy. *Health physics*, 103(5):652, 2012.
- [55] Paul B Romesser, Oren Cahlon, Eli D Scher, Eugen B Hug, Kevin Sine, Carl DeSelm, Jana L Fox, Dennis Mah, Madhur K Garg, John Han-Chih Chang, et al. Proton beam reirradiation for recurrent head and neck cancer: multi-institutional report on feasibility and early outcomes. *International Journal of Radiation Oncology* Biology* Physics*, 95(1):386–395, 2016.
- [56] Jeffrey Q Dinh, Anita Mahajan, Matthew B Palmer, and David R Grosshans. Particle therapy for central nervous system tumors in pediatric and adult patients. *Translational Cancer Research*, 1(3):137–149, 2012.
- [57] Jaeman Son, Se Byeong Lee, Youngkyung Lim, Sung Yong Park, Kwanho Cho, Myonggeun Yoon, and Dongho Shin. Development of optical fiber based measurement system for the verification of entrance dose map in pencil beam scanning proton beam. *Sensors*, 18(1):227, 2018.
- [58] Antony J Lomax, Eros Pedroni, Hans Peter Rutz, and Gudrun Goitein. The clinical potential of intensity modulated proton therapy. *Zeitschrift für Medizinische Physik*, 14(3):147–152, 2004.
- [59] Xiaodong Zhang, Yupeng Li, Xiaoning Pan, Li Xiaoqiang, Radhe Mohan, Ritsuko Komaki, James D Cox, and Joe Y Chang. Intensity-modulated proton therapy reduces the dose to normal tissue compared with intensity-

- modulated radiation therapy or passive scattering proton therapy and enables individualized radical radiotherapy for extensive stage iiib non-small-cell lung cancer: a virtual clinical study. *International Journal of Radiation Oncology* Biology* Physics*, 77(2):357–366, 2010.
- [60] Liyong Lin, Kevin Souris, Minglei Kang, Adam Glick, Haibo Lin, Sheng Huang, Kristin Stützer, Guillaume Janssens, Edmond Sterpin, John A Lee, et al. Evaluation of motion mitigation using abdominal compression in the clinical implementation of pencil beam scanning proton therapy of liver tumors. *Medical physics*, 44(2):703–712, 2017.
- [61] Stephen Dowdell, Clemens Grassberger, and Harald Paganetti. Four-dimensional monte carlo simulations demonstrating how the extent of intensity-modulation impacts motion effects in proton therapy lung treatments. *Medical physics*, 40(12):121713, 2013.
- [62] Laleh Kardar, Yupeng Li, Xiaoqiang Li, Heng Li, Wenhua Cao, Joe Y Chang, Li Liao, Ronald X Zhu, Narayan Sahoo, Michael Gillin, et al. Evaluation and mitigation of the interplay effects of intensity modulated proton therapy for lung cancer in a clinical setting. *Practical radiation oncology*, 4(6):e259–e268, 2014.
- [63] Ye Zhang, Isabel Huth, Martin Wegner, Damien Charles Weber, and Antony John Lomax. An evaluation of rescanning technique for liver tumour treatments using a commercial pbs proton therapy system. *Radiotherapy and oncology*, 121(2):281–287, 2016.
- [64] Tom Depuyt. Proton therapy technology in the clinic. 4th International Symposium on the System of Radiological Protection of ICRP and the 2nd European Radiation Protection Research Week of the European Research Platforms, 2017.
- [65] Marco Durante and Jay S Loeffler. Charged particles in radiation oncology. *Nature reviews Clinical oncology*, 7(1):37–43, 2010.
- [66] Marcel Van Herk, Peter Remeijer, Coen Rasch, and Joos V Lebesque. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. *International Journal of Radiation Oncology* Biology* Physics*, 47(4):1121–1135, 2000.
- [67] Marcel Van Herk. Errors and margins in radiotherapy. In *Seminars in radiation oncology*, volume 14, pages 52–64. Elsevier, 2004.

- [68] Eric K Hansen, M Kara Bucci, Jeanne M Quivey, Vivian Weinberg, and Ping Xia. Repeat ct imaging and replanning during the course of imrt for head-and-neck cancer. *International Journal of Radiation Oncology* Biology* Physics*, 64(2):355–362, 2006.
- [69] Eugene J Koay, David Lege, Radhe Mohan, Ritsuko Komaki, James D Cox, and Joe Y Chang. Adaptive/nonadaptive proton radiation planning and outcomes in a phase ii trial for locally advanced non-small cell lung cancer. *International Journal of Radiation Oncology* Biology* Physics*, 84(5):1093–1100, 2012.
- [70] Jagt Thyrza Z. *Automated Online Adaptive Intensity Modulated Proton Therapy*. PhD thesis, Erasmus University Rotterdam, 2020.
- [71] F Albertini, EB Hug, and Antony J Lomax. Is it necessary to plan with safety margins for actively scanned proton therapy? *Physics in Medicine & Biology*, 56(14):4399, 2011.
- [72] Albin Fredriksson, Anders Forsgren, and Björn Hårdemark. Minimax optimization for handling range and setup uncertainties in proton therapy. *Medical physics*, 38(3):1672–1684, 2011.
- [73] D Pflugfelder, JJ Wilkens, and U Oelfke. Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy. *Physics in Medicine & Biology*, 53(6):1689, 2008.
- [74] Jan Unkelbach, Thomas Bortfeld, Benjamin C Martin, and Martin Soukup. Reducing the sensitivity of impt treatment plans to setup errors and range uncertainties via probabilistic treatment planning. *Medical physics*, 36(1):149–163, 2009.
- [75] Michael Baumann, Mechthild Krause, Jens Overgaard, Jürgen Debus, Søren M Bentzen, Juliane Daartz, Christian Richter, Daniel Zips, and Thomas Bortfeld. Radiation oncology in the era of precision medicine. *Nature Reviews Cancer*, 16(4):234–249, 2016.
- [76] Lisanne V van Dijk, Roel JHM Steenbakkers, Bennie ten Haken, Hans Paul van der Laan, Aart A van 't Veld, Johannes A Langendijk, and Erik W Korevaar. Robust intensity modulated proton therapy (impt) increases estimated clinical benefit in head and neck cancer patients. *PloS one*, 11(3):e0152477, 2016.

- [77] Wei Liu, Steven J Frank, Xiaoqiang Li, Yupeng Li, Peter C Park, Lei Dong, X Ronald Zhu, and Radhe Mohan. Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers. *Med. Phys.*, 40(5):051711, May 2013.
- [78] Wei Liu, Radhe Mohan, Peter Park, Zhong Liu, Heng Li, Xiaoqiang Li, Yupeng Li, Richard Wu, Narayan Sahoo, Lei Dong, et al. Dosimetric benefits of robust treatment planning for intensity modulated proton therapy for base-of-skull cancers. *Practical radiation oncology*, 4(6):384–391, 2014.
- [79] Cássia O Ribeiro, Sabine Visser, Erik W Korevaar, Nanna M Sijtsema, R Melissa Anakotta, Margriet Dieters, Stefan Both, Johannes A Langendijk, Robin Wijsman, Christina T Muijs, et al. Towards the clinical implementation of intensity-modulated proton therapy for thoracic indications with moderate motion: Robust optimised plan evaluation by means of patient and machine specific information. *Radiotherapy and Oncology*, 157:210–218, 2021.
- [80] BK Yeh, RH Georges, XR Zhu, MB Palmer, MV Amin, JP Cheung, L Dong, J Nuanjing, JA Garcia, and SJ Frank. Adaptive replanning is required during intensity modulated proton therapy for head-and-neck cancers. *International Journal of Radiation Oncology, Biology, Physics*, 84(3):S56–S57, 2012.
- [81] Harald Paganetti, Pablo Botas, Gregory C Sharp, and Brian Winey. Adaptive proton therapy. *Phys. Med. Biol.*, 66(22), November 2021.
- [82] Yune Kwong, Alexandra Olimpia Mel, Greg Wheeler, and John M Troupis. Four-dimensional computed tomography (4dct): A review of the current status and applications. *Journal of medical imaging and radiation oncology*, 59(5):545–554, 2015.
- [83] Vincent Grégoire, Peter Levendag, Kian K. Ang, Jacques Bernier, Marijel Braaksma, Volker Budach, Cliff Chao, Emmanuel Coche, Jay S. Cooper, Guy Cosnard, Avraham Eisbruch, Samy El-Sayed, Bahman Emami, Cai Grau, Marc Hamoir, Nancy Lee, Philippe Maingon, Karin Muller, and Hervé Reyhler. Ct-based delineation of lymph node levels and related ctvs in the node-negative neck: Dahanca, eortc, gortec, ncic, rtog consensus guidelines. *Radiotherapy and Oncology*, 69(3):227–236, 2003.
- [84] Hans A Bethe and Julius Ashkin. Experimental nuclear physics. *Wiley, New York*, 1953.

- [85] P Sigmund. Particle penetration and radiation effects. series in solid-state physics, 151, 2006.
- [86] Vicki T Taasti, Christian Bäumer, Christina V Dahlgren, Amanda J Deisher, Malte Ellerbrock, Jeffrey Free, Joanna Gora, Anna Kozera, Antony J Lomax, Ludovic De Marzi, et al. Inter-centre variability of ct-based stopping-power prediction in particle therapy: survey-based evaluation. *Physics and imaging in radiation oncology*, 6:25–30, 2018.
- [87] Uwe Schneider, Eros Pedroni, and Antony Lomax. The calibration of ct hounsfield units for radiotherapy treatment planning. *Physics in Medicine & Biology*, 41(1):111, 1996.
- [88] Michael F Moyers, Milind Sardesai, Sean Sun, and Daniel W Miller. Ion stopping powers and ct numbers. *Medical Dosimetry*, 35(3):179–194, 2010.
- [89] Patrick Wohlfahrt, Christian Möhler, Christian Richter, and Steffen Greilich. 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, 2018.
- [90] Jatinder Saini, Erik Traneus, Dominic Maes, Rajesh Regmi, Stephen R Bowen, Charles Bloch, and Tony Wong. Advanced proton beam dosimetry part i: review and performance evaluation of dose calculation algorithms. *Translational lung cancer research*, 7(2):171, 2018.
- [91] Linda Hong, Michael Goitein, Marta Bucciolini, Robert Comiskey, Bernard Gottschalk, Skip Rosenthal, Chris Serago, and Marcia Urie. A pencil beam algorithm for proton dose calculations. *Physics in Medicine & Biology*, 41(8):1305, 1996.
- [92] Jefferson Sorriaux, Mauro Testa, Harald Paganetti, J Orban de Xivry, John Aldo Lee, Erik Traneus, Kevin Souris, Stefaan Vynckier, and Edmond Sterpin. Experimental assessment of proton dose calculation accuracy in inhomogeneous media. *Physica Medica*, 38:10–15, 2017.
- [93] Clemens Grassberger, Juliane Daartz, Stephen Dowdell, Thomas Ruggieri, Greg Sharp, and Harald Paganetti. Quantification of proton dose calculation accuracy in the lung. *International Journal of Radiation Oncology* Biology* Physics*, 89(2):424–430, 2014.
- [94] Pablo P Yepes, Dragan Mirkovic, and Phillip J Taddei. A gpu implementa-

- tion of a track-repeating algorithm for proton radiotherapy dose calculations. *Physics in Medicine & Biology*, 55(23):7107, 2010.
- [95] Xun Jia, Jan Schümann, Harald Paganetti, and Steve B Jiang. Gpu-based fast monte carlo dose calculation for proton therapy. *Physics in Medicine & Biology*, 57(23):7783, 2012.
- [96] Kevin Souris, John Aldo Lee, and Edmond Sterpin. Fast multipurpose monte carlo simulation for proton therapy using multi-and many-core cpu architectures. *Medical physics*, 43(4):1700–1712, 2016.
- [97] Wei Liu, Xiaodong Zhang, Yupeng Li, and Radhe Mohan. Robust optimization of intensity modulated proton therapy. *Medical physics*, 39(2):1079–1091, 2012.
- [98] Gregory Buti, Kevin Souris, Ana Maria Barragán Montero, John Aldo Lee, and Edmond Sterpin. Towards fast and robust 4d optimization for moving tumors with scanned proton therapy. *Medical physics*, 46(12):5434–5443, 2019.
- [99] Gregory Buti, Kevin Souris, Ana M Barragán Montero, Marie Cohilis, John A Lee, and Edmond Sterpin. Accelerated robust optimization algorithm for proton therapy treatment planning. *Medical Physics*, 47(7):2746–2754, 2020.
- [100] Edmond Sterpin, Sara T Rivas, F Van den Heuvel, B George, John Aldo Lee, and Kevin Souris. Development of robustness evaluation strategies for enabling statistically consistent reporting. *Physics in Medicine & Biology*, 66(4):045002, 2021.
- [101] Erik W Korevaar, Steven J M Habraken, Daniel Scandurra, Roel G J Kierkels, Mirko Unipan, Martijn G C Eenink, Roel J H Steenbakkers, Stephanie G Peeters, Jaap D Zindler, Mischa Hoogeman, and Johannes A Langendijk. Practical robustness evaluation in radiotherapy – a photon and proton-proof alternative to PTV-based plan evaluation, 2019.
- [102] G Meier, R Besson, A Nanz, S Safai, and A J Lomax. Independent dose calculations for commissioning, quality assurance and dose reconstruction of PBS proton therapy, 2015.
- [103] M Matter, Lena Nenoff, G Meier, Damien C Weber, Antony J Lomax, and Francesca Albertini. Alternatives to patient specific verification measure-

- ments in proton therapy: a comparative experimental study with intentional errors. *Physics in Medicine & Biology*, 63(20):205014, 2018.
- [104] Heng Li, Narayan Sahoo, Falk Poenisch, Kazumichi Suzuki, Yupeng Li, Xiaoliang Li, Xiaodong Zhang, Andrew K Lee, Michael T Gillin, and X Ronald Zhu. Use of treatment log files in spot scanning proton therapy as part of patient-specific quality assurance. *Medical physics*, 40(2):021703, 2013.
- [105] C-M Charlie Ma and Kamen Paskalev. In-room ct techniques for image-guided radiation therapy. *Medical dosimetry*, 31(1):30–39, 2006.
- [106] Konrad P Nesteruk, Mislav Bobić, Arthur Lalonde, Brian A Winey, Antony J Lomax, and Harald Paganetti. Ct-on-rails versus in-room cbct for on-line daily adaptive proton therapy of head-and-neck cancers. *Cancers*, 13(23):5991, 2021.
- [107] Laura Ann Rechner, Maja Vestmø Maraldo, Ivan Richter Vogelius, Xiaorong Ronald Zhu, Bouthaina Shbib Dabaja, Nils Patrik Brodin, Peter Meidahl Petersen, Lena Specht, and Marianne Camille Aznar. Life years lost attributable to late effects after radiotherapy for early stage hodgkin lymphoma: The impact of proton therapy and/or deep inspiration breath hold. *Radiotherapy and Oncology*, 125(1):41–47, 2017.
- [108] Amy C Moreno, Jillian R Gunther, Sarah Milgrom, C David Fuller, Tyler Williamson, Amy Liu, Richard Wu, X Ronald Zhu, Bouthaina S Dabaja, and Chelsea C Pinnix. Effect of deep inspiration breath hold on normal tissue sparing with intensity modulated radiation therapy versus proton therapy for mediastinal lymphoma. *Advances in Radiation Oncology*, 5(6):1255–1266, 2020.
- [109] Geneviève Van Ooteghem, Damien Dasnoy-Sumell, John Aldo Lee, and Xavier Geets. Mechanically-assisted and non-invasive ventilation for radiation therapy: a safe technique to regularize and modulate internal tumour motion. *Radiotherapy and Oncology*, 141:283–291, 2019.
- [110] L Vander Veken, G Van Ooteghem, B Ghaye, A Razavi, and X Geets. Deep inspiration breath hold (dibh) for lung sabr under mechanically-assisted and non-invasive ventilation (maniv). In *MEDICAL PHYSICS*, volume 49, pages E269–E269. WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, 2022.
- [111] Christian Graeff, Anna Constantinescu, Robert Luchtenborg, Marco Du-

- rante, and Christoph Bert. Multigating, a 4d optimized beam tracking in scanned ion beam therapy. *Technology in cancer research & treatment*, 13(6):497–504, 2014.
- [112] Naoki Miyamoto, Kouhei Yokokawa, Seishin Takao, Taeko Matsuura, So-dai Tanaka, Shinichi Shimizu, Hiroki Shirato, and Kikuo Umegaki. Dynamic gating window technique for the reduction of dosimetric error in respiratory-gated spot-scanning particle therapy: An initial phantom study using patient tumor trajectory data. *Journal of Applied Clinical Medical Physics*, 21(4):13–21, 2020.
- [113] Antje-Christin Knopf and Antony Lomax. In vivo proton range verification: a review. *Physics in Medicine & Biology*, 58(15):R131, 2013.
- [114] Uwe Schneider and Eros Pedroni. Proton radiography as a tool for quality control in proton therapy. *Medical physics*, 22(4):353–363, 1995.
- [115] Christian Richter, Guntram Pausch, Steffen Barczyk, Marlen Priegnitz, Isabell Keitz, Julia Thiele, Julien Smeets, Francois Vander Stappen, Luca Bombelli, Carlo Fiorini, et al. First clinical application of a prompt gamma based in vivo proton range verification system. *Radiotherapy and Oncology*, 118(2):232–237, 2016.
- [116] Jon Cacicedo, Francisco Casquero, Lorea Martinez-Indart, Olga Del Hoyo, Alfonso Gomez de Iturriaga, Arturo Navarro, and Pedro Bilbao. A prospective analysis of factors that influence weight loss in patients undergoing radiotherapy. *Chinese journal of cancer*, 33(4):204, 2014.
- [117] Suzanne van Beek, Marcel Jonker, Olga Hamming-Vrieze, Abraham Al-Mamgani, Arash Navran, Peter Remeijer, and Jeroen B van de Kamer. Protocolised way to cope with anatomical changes in head & neck cancer during the course of radiotherapy. *Technical Innovations & Patient Support in Radiation Oncology*, 12:34–40, 2019.
- [118] M Palmer, T Jones, M Waddell, S Varghese, Y Zhang, M Amin, BK Yeh, R Zhu, L Dong, and SJ Frank. The optimal timing for off-line adaptive planning for bilateral head-and-neck impt is week 4. *International Journal of Radiation Oncology, Biology, Physics*, 84(3):S479–S480, 2012.
- [119] L E Henke, J A Contreras, O L Green, B Cai, H Kim, M C Roach, J R Olsen, B Fischer-Valuck, D F Mullen, R Kashani, M A Thomas, J Huang, I Zoberi, D Yang, V Rodriguez, J D Bradley, C G Robinson, P Parikh, S Mutic, and

- J Michalski. Magnetic resonance Image-Guided radiotherapy (MRIgRT): A 4.5-year clinical experience. *Clin. Oncol.*, 30(11):720–727, November 2018.
- [120] Jan-Jakob Sonke, Joos Lebesque, and Marcel van Herk. Variability of Four-Dimensional computed tomography patient models, 2008.
- [121] Salim Balik, Elisabeth Weiss, Nuzhat Jan, Nicholas Roman, William C Sleeman, Mirek Fatyga, Gary E Christensen, Cheng Zhang, Martin J Murphy, Jun Lu, Paul Keall, Jeffrey F Williamson, and Geoffrey D Hugo. Evaluation of 4-dimensional computed tomography to 4-dimensional cone-beam computed tomography deformable image registration for lung cancer adaptive radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 86(2):372–379, June 2013.
- [122] Geoffrey Hugo, Carlos Vargas, Jian Liang, Larry Kestin, John W Wong, and Di Yan. Changes in the respiratory pattern during radiotherapy for cancer in the lung. *Radiother. Oncol.*, 78(3):326–331, March 2006.
- [123] Daniel Scandurra, Tineke WH Meijer, Jeffrey Free, Johanna GM van den Hoek, Lotta Kelder, Edwin Oldehinkel, Roel JHM Steenbakkers, Stefan Both, and Johannes A Langendijk. Evaluation of robustly optimised intensity modulated proton therapy for nasopharyngeal carcinoma. *Radiotherapy and Oncology*, 168:221–228, 2022.
- [124] Steven van de Water, Francesca Albertini, Damien C Weber, Ben J M Heijmen, Mischa S Hoogeman, and Antony J Lomax. 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(2):025020, January 2018.
- [125] Zhiyong Yang, Xiaodong Zhang, Xianliang Wang, X Ronald Zhu, Brandon Gunn, Steven J Frank, Yu Chang, Qin Li, Kunyu Yang, Gang Wu, Li Liao, Yupeng Li, Mei Chen, and Heng Li. Multiple-CT optimization: An adaptive optimization method to account for anatomical changes in intensity-modulated proton therapy for head and neck cancers. *Radiother. Oncol.*, 142:124–132, January 2020.
- [126] Arthur Lalonde, Mislav Bobić, Brian Winey, Joost Verburg, Gregory C Sharp, and Harald Paganetti. Anatomic changes in head and neck intensity-modulated proton therapy: Comparison between robust optimization and online adaptation. *Radiother. Oncol.*, 159:39–47, June 2021.

- [127] Xianliang Wang, Heng Li, Xiaorong Ronald Zhu, Qing Hou, Li Liao, Bo Jiang, Yupeng Li, Pei Wang, Jinyi Lang, and Xiaodong Zhang. Multiple-ct optimization of intensity-modulated proton therapy—is it possible to eliminate adaptive planning? *Radiotherapy and Oncology*, 128(1):167–173, 2018.
- [128] Yoshikazu Maeda, Yoshitaka Sato, Hiroki Minami, Yutaka Yasukawa, Kazutaka Yamamoto, Hiroyasu Tamamura, Satoshi Shibata, Sayuri Bou, Makoto Sasaki, Yuji Tameshige, 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):1832–1843, 2018.
- [129] Kubeš Jiří, Vondráček Vladimír, Andrlík Michal, Navrátil Matěj, Sláviková Silvia, Vítek Pavel, Dědečková Kateřina, Prausová Jana, Ondrová Barbora, Rotnáglová Eliška, et al. Proton pencil-beam scanning radiotherapy in the treatment of nasopharyngeal cancer: dosimetric parameters and 2-year results. *European Archives of Oto-Rhino-Laryngology*, 278(3):763–769, 2021.
- [130] Hans Paul van der Laan, R Melissa Anakotta, Erik W Korevaar, Margriet Dieters, J Fred Ubbels, Robin Wijsman, Nanna M Sijtsema, Stefan Both, Johannes A Langendijk, Christina T Muijs, et al. Organ sparing potential and inter-fraction robustness of adaptive intensity modulated proton therapy for lung cancer. *Acta Oncologica*, 58(12):1775–1782, 2019.
- [131] Jasper Nijkamp, Rianne de Jong, Jan-Jakob Sonke, Peter Remeijer, Corine van Vliet, and Corrie Marijnen. Target volume shape variation during hypofractionated preoperative irradiation of rectal cancer patients. *Radiother. Oncol.*, 92(2):202–209, August 2009.
- [132] Curtis Woodford, Slav Yartsev, A Rashid Dar, Glenn Bauman, and Jake Van Dyk. Adaptive radiotherapy planning on decreasing gross tumor volumes as seen on megavoltage computed tomography images. *Int. J. Radiat. Oncol. Biol. Phys.*, 69(4):1316–1322, November 2007.
- [133] Patrick A Kupelian, Chester Ramsey, Sanford L Meeks, Twyla R Willoughby, Alan Forbes, Thomas H Wagner, and Katja M Langen. Serial megavoltage CT imaging during external beam radiotherapy for non-small-cell lung cancer: observations on tumor regression during treatment. *Int. J. Radiat. Oncol. Biol. Phys.*, 63(4):1024–1028, November 2005.
- [134] CE Shelley, Lisa H Barraclough, Claire L Nelder, SJ Otter, and AJ Stewart. Adaptive radiotherapy in the management of cervical cancer: review of

- strategies and clinical implementation. *Clinical Oncology*, 33(9):579–590, 2021.
- [135] Francesca Albertini, Michael Matter, Lena Nenoff, Ye Zhang, and Antony Lomax. Online daily adaptive proton therapy. *Br. J. Radiol.*, 93(1107):20190594, March 2020.
- [136] Dennis Winkel, Gijsbert H Bol, Petra S Kroon, Bram van Asselen, Sara S Hackett, Anita M Werensteijn-Honingh, Martijn P W Intven, Wietse S C Eppinga, Rob H N Tijssen, Linda G W Kerkmeijer, Hans C J de Boer, Stella Mook, Gert J Meijer, Jochem Hes, Mirjam Willemsen-Bosman, Eline N de Groot-van Breugel, Ina M Jürgenliemk-Schulz, and Bas W Raaymakers. Adaptive radiotherapy: The elekta unity MR-linac concept. *Clin Transl Radiat Oncol*, 18:54–59, September 2019.
- [137] Mikel Byrne, Ben Archibald-Heeren, Yunfei Hu, Amy Teh, Rhea Beserminji, Emma Cai, Guilin Liu, Angela Yates, James Rijken, Nick Collett, and Trent Aland. Varian ethos online adaptive radiotherapy for prostate cancer: Early results of contouring accuracy, treatment plan quality, and treatment time. *J. Appl. Clin. Med. Phys.*, 23(1):e13479, January 2022.
- [138] Carri K. Glide-Hurst, Percy Lee, Adam D. Yock, Jeffrey R. Olsen, Minsong Cao, Farzan Siddiqui, William Parker, Anthony Doemer, Yi Rong, Amar U. Kishan, Stanley H. Benedict, X. Allen Li, Beth A. Erickson, Jason W. Sohn, Ying Xiao, and Evan Wuthrick. Adaptive radiation therapy (art) strategies and technical considerations: A state of the art review from nrg oncology. *International Journal of Radiation Oncology*Biology*Physics*, 109(4):1054–1075, 2021.
- [139] Konrad P Nesteruk, Mislav Bobić, Gregory C Sharp, Arthur Lalonde, Brian A Winey, Lena Nenoff, Antony J Lomax, and Harald Paganetti. Low-dose computed tomography scanning protocols for online adaptive proton therapy of head-and-neck cancers. *Cancers*, 14(20):5155, 2022.
- [140] Yang-Kyun Park, Gregory C Sharp, Justin Phillips, and Brian A Winey. Proton dose calculation on scatter-corrected CBCT image: Feasibility study for adaptive proton therapy. *Med. Phys.*, 42(8):4449–4459, August 2015.
- [141] Kazuhiro Arai, Noriyuki Kadoya, Takahiro Kato, Hiromitsu Endo, Shinya Komori, Yoshitomo Abe, Tatsuya Nakamura, Hitoshi Wada, Yasuhiro Kikuchi, Yoshihiro Takai, and Keiichi Jingu. Feasibility of CBCT-based proton dose calculation using a histogram-matching algorithm in proton beam

therapy, 2017.

- [142] Christopher Kurz, Florian Kamp, Yang-Kyun Park, Christoph Zöllner, Simon Rit, David Hansen, Mark Podesta, Gregory C Sharp, Minglun Li, Michael Reiner, Jan Hofmaier, Sebastian Nepl, Christian Thieke, Reinoud Nijhuis, Ute Ganswindt, Claus Belka, Brian A Winey, Katia Parodi, and Guillaume Landry. Investigating deformable image registration and scatter correction for CBCT-based dose calculation in adaptive IMPT, 2016.
- [143] Adrian Thummerer, Paolo Zaffino, Arturs Meijers, Gabriel Guterres Marmitt, Joao Seco, Roel J H M Steenbakkers, Johannes A Langendijk, Stefan Both, Maria F Spadea, and Antje C Knopf. Comparison of CBCT based synthetic CT methods suitable for proton dose calculations in adaptive proton therapy. *Phys. Med. Biol.*, 65(9):095002, April 2020.
- [144] Sinan Irmak, Dietmar Georg, and Wolfgang Lechner. Comparison of CBCT conversion methods for dose calculation in the head and neck region. *Z. Med. Phys.*, 30(4):289–299, November 2020.
- [145] Allen M Chen, Megan E Daly, Jing Cui, Mathew Mathai, Stanley Benedict, and James A Purdy. Clinical outcomes among patients with head and neck cancer treated by intensity-modulated radiotherapy with and without adaptive replanning. *Head & neck*, 36(11):1541–1546, 2014.
- [146] Jolien Heukelom and Clifton David Fuller. Head and neck cancer adaptive radiation therapy (ART): Conceptual considerations for the informed clinician. *Semin. Radiat. Oncol.*, 29(3):258–273, July 2019.
- [147] Karen Lim, James Stewart, Valerie Kelly, Jason Xie, Kristy K Brock, Joanne Moseley, Young-Bin Cho, Anthony Fyles, Anna Lundin, Henrik Rehbinder, Johan Löf, David A Jaffray, and Michael Milosevic. Dosimetrically triggered adaptive intensity modulated radiation therapy for cervical cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, 90(1):147–154, September 2014.
- [148] Christian Dial, Elisabeth Weiss, Jeffrey V Siebers, and Geoffrey D Hugo. Benefits of adaptive radiation therapy in lung cancer as a function of replanning frequency. *Med. Phys.*, 43(4):1787, April 2016.
- [149] A Meijers, A Jakobi, K Stützer, G Guterres Marmitt, S Both, J A Langendijk, C Richter, and A Knopf. Log file-based dose reconstruction and accumulation for 4D adaptive pencil beam scanned proton therapy in a clinical treatment planning system: Implementation and proof-of-concept,

2019.

- [150] Dennis Mackin, Yupeng Li, Michael B Taylor, Matthew Kerr, Charles Holmes, Narayan Sahoo, Falk Poenisch, Heng Li, Jim Lii, Richard Amos, Richard Wu, Kazumichi Suzuki, Michael T Gillin, X Ronald Zhu, and Xiaodong Zhang. Improving spot-scanning proton therapy patient specific quality assurance with HPlusQA, a second-check dose calculation engine, 2013.
- [151] X Ronald Zhu, Yupeng Li, Dennis Mackin, Heng Li, Falk Poenisch, Andrew K Lee, Anita Mahajan, Steven J Frank, Michael T Gillin, Narayan Sahoo, and Xiaodong Zhang. Towards effective and efficient patient-specific quality assurance for spot scanning proton therapy. *Cancers*, 7(2):631–647, April 2015.
- [152] Bin Cai, Olga L Green, Rojano Kashani, Vivian L Rodriguez, Sasa Mutic, and Deshan Yang. A practical implementation of physics quality assurance for photon adaptive radiotherapy. *Z. Med. Phys.*, 28(3):211–223, August 2018.
- [153] Chunbo Liu, Meng Wei Ho, Jiyeon Park, Wen Chien Hsi, Xiaoying Liang, Zuofeng Li, Yuntao Song, Hansheng Feng, and Yawei Zhang. Fast MCsquare-Based independent dose verification platform for pencil beam scanning proton therapy, 2021.
- [154] Simona Toscano, Kevin Souris, Carles Goma, A Barragán-Montero, S Puydupin, F Vander Stappen, Guillaume Janssens, A Matic, Xavier Geets, and Edmond Sterpin. Impact of machine log-files uncertainties on the quality assurance of proton pencil beam scanning treatment delivery. *Physics in Medicine & Biology*, 64(9):095021, 2019.
- [155] Olga L Green, Lauren E Henke, and Geoffrey D Hugo. Practical clinical workflows for online and offline adaptive radiation therapy. *Semin. Radiat. Oncol.*, 29(3):219–227, July 2019.
- [156] Annette Boejen, Anne Vestergaard, Lone Hoffmann, Mai-Britt Ellegaard, Anne-Mette Rasmussen, Ditte Møller, Ludvig P Muren, and Cai Grau. A learning programme qualifying radiation therapists to manage daily online adaptive radiotherapy. *Acta Oncol.*, 54(9):1697–1701, July 2015.
- [157] H A McNair, S Hafeez, H Taylor, S Lalondrelle, F McDonald, V N Hansen, and R Huddart. Radiographer-led plan selection for bladder cancer radio-

- therapy: initiating a training programme and maintaining competency. *Br. J. Radiol.*, 88(1048):20140690, April 2015.
- [158] Indrin J Chetty and Mihaela Rosu-Bubulac. Deformable registration for dose accumulation, 2019.
- [159] Chang Liu, Jinkoo Kim, Akila Kumarasiri, Essa Mayyas, Stephen L Brown, Ning Wen, Farzan Siddiqui, and Indrin J Chetty. An automated dose tracking system for adaptive radiation therapy. *Comput. Methods Programs Biomed.*, 154:1–8, February 2018.
- [160] Jenny Bertholet, Gail Anastasi, David Noble, Arjan Bel, Ruud van Leeuwen, Toon Roggen, Michael Duchateau, Sara Pilskog, Cristina Garibaldi, Nina Tilly, Rafael García-Mollá, Jorge Bonaque, Uwe Oelfke, Marianne C Aznar, and Ben Heijmen. Patterns of practice for adaptive and real-time radiation therapy (POP-ART RT) part II: Offline and online plan adaption for interfractional changes. *Radiother. Oncol.*, 153:88–96, December 2020.
- [161] Ana Barragán-Montero, Umair Javaid, Gilmer Valdés, Dan Nguyen, Paul Desbordes, Benoit Macq, Siri Willems, Liesbeth Vandewinckele, Mats Holmström, Fredrik Löfman, et al. Artificial intelligence and machine learning for medical imaging: A technology review. *Physica Medica*, 83:242–256, 2021.
- [162] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [163] Joseph Harms, Yang Lei, Tonghe Wang, Mark McDonald, Beth Ghavidel, William Stokes, Walter J Curran, Jun Zhou, Tian Liu, and Xiaofeng Yang. Cone-beam CT-derived relative stopping power map generation via deep learning for proton radiotherapy. *Med. Phys.*, 47(9):4416–4427, September 2020.
- [164] Satoshi Kida, Shizuo Kaji, Kanabu Nawa, Toshikazu Imae, Takahiro Nakamoto, Sho Ozaki, Takeshi Ohta, Yuki Nozawa, and Keiichi Nakagawa. Visual enhancement of cone-beam CT by use of CycleGAN. *Med. Phys.*, 47(3):998–1010, March 2020.
- [165] Yang Lei, Joseph Harms, Tonghe Wang, Yingzi Liu, Hui-Kuo Shu, Ashesh B Jani, Walter J Curran, Hui Mao, Tian Liu, and Xiaofeng Yang. MRI-only based synthetic CT generation using dense cycle consistent generative

- adversarial networks. *Med. Phys.*, 46(8):3565–3581, August 2019.
- [166] Xiao Liang, Liyuan Chen, Dan Nguyen, Zhiguo Zhou, Xuejun Gu, Ming Yang, Jing Wang, and Steve Jiang. Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. *Phys. Med. Biol.*, 64(12):125002, June 2019.
- [167] Yingzi Liu, Yang Lei, Tonghe Wang, Yabo Fu, Xiangyang Tang, Walter J Curran, Tian Liu, Pretesh Patel, and Xiaofeng Yang. CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy. *Med. Phys.*, 47(6):2472–2483, June 2020.
- [168] Matteo Maspero, Antonetta C Houweling, Mark H F Savenije, Tristan C F van Heijst, Joost J C Verhoeff, Alexis N T J Kotte, and Cornelis A T van den Berg. A single neural network for cone-beam computed tomography-based radiotherapy of head-and-neck, lung and breast cancer. *Phys Imaging Radiat Oncol*, 14:24–31, April 2020.
- [169] Liyuan Chen, Xiao Liang, Chenyang Shen, Steve Jiang, and Jing Wang. Synthetic CT generation from CBCT images via deep learning. *Med. Phys.*, 47(3):1115–1125, March 2020.
- [170] Arthur Lalonde, Brian Winey, Joost Verburg, Harald Paganetti, and Gregory C Sharp. Evaluation of CBCT scatter correction using deep convolutional neural networks for head and neck adaptive proton therapy. *Phys. Med. Biol.*, 65(24), December 2020.
- [171] Adrian Thummerer, Bas A de Jong, Paolo Zaffino, Arturs Meijers, Gabriel Guterres Marmitt, Joao Seco, Roel J H M Steenbakkers, Johannes A Langendijk, Stefan Both, Maria F Spadea, and Antje C Knopf. Comparison of the suitability of CBCT- and MR-based synthetic CTs for daily adaptive proton therapy in head and neck patients. *Phys. Med. Biol.*, 65(23):235036, December 2020.
- [172] Tim Lustberg, Johan van Soest, Mark Gooding, Devis Peressutti, Paul Aljabar, Judith van der Stoep, Wouter van Elmpt, and Andre Dekker. Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer. *Radiotherapy and Oncology*, 126(2):312–317, 2018.
- [173] Humera Shaziya, K Shyamala, and Raniah Zaheer. Automatic lung segmentation on thoracic ct scans using u-net convolutional network. In *2018*

- International conference on communication and signal processing (ICCSP)*, pages 0643–0647. IEEE, 2018.
- [174] Bulat Ibragimov and Lei Xing. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. *Med. Phys.*, 44(2):547–557, February 2017.
 - [175] Yngve Mardal Moe, Aurora Rosvoll Groendahl, Martine Mulstad, Oliver Tomic, Ulf Indahl, Einar Dale, Eirik Malinen, and Cecilia Marie Futsaether. Deep learning for automatic tumour segmentation in pet/ct images of patients with head and neck cancers. *arXiv preprint arXiv:1908.00841*, 2019.
 - [176] Shihao Li, Jianghong Xiao, Ling He, Xingchen Peng, and Xuedong Yuan. The tumor target segmentation of nasopharyngeal cancer in ct images based on deep learning methods. *Technology in cancer research & treatment*, 18:1533033819884561, 2019.
 - [177] Adrian L Breto, Benjamin Spieler, Olmo Zavala-Romero, Mohammad Al-husseini, Nirav V Patel, David A Asher, Isaac R Xu, Jacqueline B Baikovitz, Eric A Mellon, John C Ford, Radka Stoyanova, and Lorraine Portelance. Deep learning for Per-Fraction automatic segmentation of gross tumor volume (GTV) and organs at risk (OARs) in adaptive radiotherapy of cervical cancer. *Front. Oncol.*, 12:854349, May 2022.
 - [178] Zhikai Liu, Fangjie Liu, Wanqi Chen, Yinjie Tao, Xia Liu, Fuquan Zhang, Jing Shen, Hui Guan, Hongnan Zhen, Shaobin Wang, Qi Chen, Yu Chen, and Xiaorong Hou. Automatic segmentation of clinical target volume and Organs-at-Risk for breast conservative radiotherapy using a convolutional neural network. *Cancer Manag. Res.*, 13:8209–8217, November 2021.
 - [179] S A Yoganathan, Siji Nojin Paul, Satheesh Paloor, Tarraf Torfeh, Suparna Halsnad Chandramouli, Rabih Hammoud, and Noora Al-Hammadi. Automatic segmentation of magnetic resonance images for high-dose-rate cervical cancer brachytherapy using deep learning. *Med. Phys.*, 49(3):1571–1584, March 2022.
 - [180] J van der Veen, S Willems, H Bollen, F Maes, and S Nuyts. Deep learning for elective neck delineation: More consistent and time efficient. *Radiother. Oncol.*, 153:180–188, December 2020.
 - [181] Charlotte L Brouwer, Anna M Dinkla, Liesbeth Vandewinckele, Wouter Crijs, Michaël Claessens, Dirk Verellen, and Wouter van Elmpt. Machine

- learning applications in radiation oncology: Current use and needs to support clinical implementation. *Physics and imaging in radiation oncology*, 16:144–148, 2020.
- [182] Siri Willems, Wouter Crijns, Agustina La Greca Saint-Estevan, Julie Van Der Veen, David Robben, Tom Depuydt, Sandra Nuyts, Karin Haustermans, and Frederik Maes. Clinical implementation of DeepVoxNet for Auto-Delineation of organs at risk in head and neck cancer patients in radiotherapy, 2018.
- [183] J Van der Veen, S Willems, S Deschuymer, D Robben, W Crijns, F Maes, and S Nuyts. Benefits of deep learning for delineation of organs at risk in head and neck cancer. *Radiotherapy and Oncology*, 138:68–74, 2019.
- [184] Shadab Momin, Yabo Fu, Yang Lei, Justin Roper, Jeffrey D Bradley, Walter J Curran, Tian Liu, and Xiaofeng Yang. Knowledge-based radiation treatment planning: A data-driven method survey. *Journal of Applied Clinical Medical Physics*, 22(8):16–44, 2021.
- [185] Ana Barragán-Montero, Adrien Bibal, Margerie Huet Dastarac, Camille Dragnet, Gilmer Valdés, Dan Nguyen, Siri Willems, Liesbeth Vandewinckele, Mats Holmström, Fredrik Löfman, Kevin Souris, Edmond Sterpin, and John A Lee. Towards a safe and efficient clinical implementation of machine learning in radiation oncology by exploring model interpretability, explainability and data-model dependency. *Phys. Med. Biol.*, 67(11), May 2022.
- [186] Tucker J Netherton, Carlos E Cardenas, Dong Joo Rhee, Laurence E Court, and Beth M Beadle. The emergence of artificial intelligence within radiation oncology treatment planning. *Oncology*, 99(2):124–134, 2021.
- [187] Daniel Jarrett, Eleanor Stride, Katherine Vallis, and Mark J Gooding. Applications and limitations of machine learning in radiation oncology. *Br. J. Radiol.*, 92(1100), August 2019.
- [188] Xun Jia, Peter Ziegenhein, and Steve B Jiang. GPU-based high-performance computing for radiation therapy, 2014.
- [189] Myles Spalding, Anthony Walsh, and Trent Aland. Evaluation of a new GPU-enabled VMAT multi-criteria optimisation plan generation algorithm. *Med. Dosim.*, 45(4):368–373, July 2020.
- [190] Trent Aland, Anthony Walsh, Mark Jones, Angelina Piccini, and Aimee

- Devlin. Accuracy and efficiency of graphics processing unit (GPU) based acuros XB dose calculation within the varian eclipse treatment planning system. *Med. Dosim.*, 44(3):219–225, 2019.
- [191] MPW Intven, SR de Mol van Otterloo, S Mook, PAH Doornaert, EN de Groot-van Breugel, GG Sikkes, ME Willemsen-Bosman, HM van Zijp, and RHN Tijssen. Online adaptive mr-guided radiotherapy for rectal cancer; feasibility of the workflow on a 1.5 t mr-linac: clinical implementation and initial experience. *Radiotherapy and Oncology*, 154:172–178, 2021.
- [192] Lukas Nierer, Chukwuka Eze, Vanessa da Silva Mendes, Juliane Braun, Patrick Thum, Rieke von Bestenbostel, Christopher Kurz, Guillaume Landry, Michael Reiner, Maximilian Niyazi, et al. Dosimetric benefit of mr-guided online adaptive radiotherapy in different tumor entities: liver, lung, abdominal lymph nodes, pancreas and prostate. *Radiation Oncology*, 17(1):1–14, 2022.
- [193] Rasmus Lübeck Christiansen, Lars Dysager, Christian Rønn Hansen, Henrik Robenhagen Jensen, Tine Schytte, Christina Junker Nyborg, Anders Smedegaard Bertelsen, Søren Nielsen Agergaard, Faisal Mahmood, Steinbjørn Hansen, et al. Online adaptive radiotherapy potentially reduces toxicity for high-risk prostate cancer treatment. *Radiotherapy and Oncology*, 167:165–171, 2022.
- [194] Patrik Sibolt, Lina M Andersson, Lucie Calmels, David Sjöström, Ulf Bjelkengren, Poul Geertsens, and Claus F Behrens. Clinical implementation of artificial intelligence-driven cone-beam computed tomography-guided online adaptive radiotherapy in the pelvic region. *Phys Imaging Radiat Oncol*, 17:1–7, January 2021.
- [195] Suk Whan Yoon, Hui Lin, Michelle Alonso-Basanta, Nate Anderson, Ontida Apinorasethkul, Karima Cooper, Lei Dong, Brian Kempsey, Jaclyn Marcel, James Metz, Ryan Scheuermann, and Taoran Li. Initial evaluation of a novel Cone-Beam CT-Based Semi-Automated online adaptive radiotherapy system for head and neck cancer treatment – a timing and automation quality study, 2020.
- [196] Lisanne GM Zwart, Francisca Ong, A Liselotte, Erik B van Dieren, Siete A Koch, Anand Bhawanie, Elisabeth de Wit, and Judith J Dasselaar. Cone-beam computed tomography-guided online adaptive radiotherapy is feasible for prostate cancer patients. *Physics and Imaging in Radiation Oncology*,

22:98–103, 2022.

- [197] Michael Matter, Lena Nenoff, Gabriel Meier, Damien C Weber, Antony J Lomax, and Francesca Albertini. Intensity modulated proton therapy plan generation in under ten seconds. *Acta oncologica*, 58(10):1435–1439, 2019.
- [198] Lena Nenoff, Michael Matter, Johanna Hedlund Lindmar, Damien Charles Weber, Antony John Lomax, and Francesca Albertini. Daily adaptive proton therapy—the key to innovative planning approaches for paranasal cancer treatments. *Acta oncologica*, 58(10):1423–1428, 2019.
- [199] P Botas, J Kim, B Winey, and H Paganetti. Online adaption approaches for intensity modulated proton therapy for head and neck patients based on cone beam CTs and monte carlo simulations. *Phys. Med. Biol.*, 64(1):015004, December 2018.
- [200] Mislav Bobić, Arthur Lalonde, Gregory C Sharp, Clemens Grassberger, Joost M Verburg, Brian A Winey, Antony John Lomax, and Harald Paganetti. Comparison of weekly and daily online adaptation for head and neck intensity-modulated proton therapy. *Phys. Med. Biol.*, January 2021.
- [201] Thyrza Jagt, Sebastiaan Breedveld, Rens Van Haveren, Ben Heijmen, and Mischa Hoogeman. An automated planning strategy for near real-time adaptive proton therapy in prostate cancer. *Physics in Medicine & Biology*, 63(13):135017, 2018.
- [202] Thyrza Z Jagt, Sebastiaan Breedveld, Rens Van Haveren, Remi A Nout, Eleftheria Astreinidou, Ben JM Heijmen, and Mischa S Hoogeman. Plan-library supported automated replanning for online-adaptive intensity-modulated proton therapy of cervical cancer. *Acta Oncologica*, 58(10):1440–1445, 2019.
- [203] Michelle Oud, Sebastiaan Breedveld, Marta Giżyńska, Michiel Kroesen, Stefan Hutschemaekers, Steven Habraken, Steven Petit, Zoltán Perkó, Ben Heijmen, and Mischa Hoogeman. An online adaptive plan library approach for intensity modulated proton therapy for head and neck cancer. *Radiotherapy and Oncology*, 176:68–75, 2022.
- [204] Aparna H Kesarwala, Christine J Ko, Holly Ning, Eric Xanthopoulos, Karl E Haglund, William P O'Meara, Charles B Simone II, and Ramesh Rengan. Intensity-modulated proton therapy for elective nodal irradiation and involved-field radiation in the definitive treatment of locally advanced non–

- small-cell lung cancer: a dosimetric study. *Clinical lung cancer*, 16(3):237–244, 2015.
- [205] Matthew J Ferris, Katherine S Martin, Jeffrey M Switchenko, Oluwatosin A Kayode, Jonathan Wolf, Quang Dang, Robert H Press, Walter J Curran, and Kristin A Higgins. Sparing cardiac substructures with optimized volumetric modulated arc therapy and intensity modulated proton therapy in thoracic radiation for locally advanced non-small cell lung cancer. *Practical radiation oncology*, 9(5):e473–e481, 2019.
- [206] R Charles Nichols, Soon N Huh, Randal H Henderson, Nancy P Mendenhall, Stella Flampouri, Zuofeng Li, Harry J D’Agostino, J Davis Cury, Dat C Pham, and Bradford S Hoppe. Proton radiation therapy offers reduced normal lung and bone marrow exposure for patients receiving dose-escalated radiation therapy for unresectable stage iii non-small-cell lung cancer: a dosimetric study. *Clinical lung cancer*, 12(4):252–257, 2011.
- [207] Joe Y Chang, Wencheng Zhang, Ritsuko Komaki, Noah C Choi, Shen Chan, Daniel Gomez, Michael O’Reilly, Melenda Jeter, Michael Gillin, Xiaorong Zhu, et al. Long-term outcome of phase i/ii prospective study of dose-escalated proton therapy for early-stage non-small cell lung cancer. *Radiotherapy and Oncology*, 122(2):274–280, 2017.
- [208] JC Ho, Heng Li, PK Allen, Xiaodong Zhang, Zhongxing Liao, X Ronald Zhu, DR Gomez, SH Lin, MT Gillin, RU Komaki, et al. Clinical outcome of intensity modulated proton therapy for non-small cell lung cancer. *International Journal of Radiation Oncology, Biology, Physics*, 93(3):S188, 2015.
- [209] Chiyoko Makita, Tatsuya Nakamura, Akinori Takada, Kanako Takayama, Motohisa Suzuki, Yusuke Azami, Takahiro Kato, Iwao Tsukiyama, Masato Hareyama, Yasuhiro Kikuchi, et al. High-dose proton beam therapy for stage i non-small cell lung cancer: Clinical outcomes and prognostic factors. *Acta Oncologica*, 54(3):307–314, 2015.
- [210] Adnan Elhammali, Pierre Blanchard, Alison Yoder, Zhongxing Liao, Xiadong Zhang, X Ronald Zhu, Pamela K Allen, Melenda Jeter, James Welsh, and Quynh-Nhu Nguyen. Clinical outcomes after intensity-modulated proton therapy with concurrent chemotherapy for inoperable non-small cell lung cancer. *Radiotherapy and Oncology*, 136:136–142, 2019.
- [211] Sebastian Van Der Voort, Steven Van De Water, Zoltán Perkó, Ben Heij-

- men, 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.
- [212] Heng Li, Xiaodong Zhang, Peter Park, Wei Liu, Joe Chang, Zhongxing Liao, Steve Frank, Yupeng Li, Falk Poenisch, Radhe Mohan, et al. Robust optimization in intensity-modulated proton therapy to account for anatomy changes in lung cancer patients. *Radiotherapy and Oncology*, 114(3):367–372, 2015.
- [213] Clemens Grassberger, Stephen Dowdell, Antony Lomax, Greg Sharp, James Shackelford, Noah Choi, Henning Willers, and Harald Paganetti. Motion interplay as a function of patient parameters and spot size in spot scanning proton therapy for lung cancer. *International Journal of Radiation Oncology* Biology* Physics*, 86(2):380–386, 2013.
- [214] Kevin L Moore. Automated radiotherapy treatment planning. In *Seminars in radiation oncology*, volume 29, pages 209–218. Elsevier, 2019.
- [215] Dirk De Ruyscher, Corinne Faivre-Finn, Ursula Nestle, Coen W Hurkmans, Cécile Le Péchoux, Allan Price, and Suresh Senan. European organisation for research and treatment of cancer recommendations for planning and delivery of high-dose, high-precision radiotherapy for lung cancer. *Journal of clinical oncology*, 28(36):5301–5310, 2010.
- [216] Tatsuya Inoue, Joachim Widder, Lisanne V van Dijk, Hideki Takegawa, Masahiko Koizumi, Masaaki Takashina, Keisuke Usui, Chie Kurokawa, Satoru Sugimoto, Anneyuko I Saito, et al. Limited impact of setup and range uncertainties, breathing motion, and interplay effects in robustly optimized intensity modulated proton therapy for stage iii non-small cell lung cancer. *International Journal of Radiation Oncology* Biology* Physics*, 96(3):661–669, 2016.
- [217] Wei Liu, Steven E Schild, Joe Y Chang, Zhongxing Liao, Yu-Hui Chang, Zhifei Wen, Jiajian Shen, Joshua B Stoker, Xiaoning Ding, Yanle Hu, et al. Exploratory study of 4d versus 3d robust optimization in intensity modulated proton therapy for lung cancer. *International Journal of Radiation Oncology* Biology* Physics*, 95(1):523–533, 2016.
- [218] Marie Wanet, Edmond Sterpin, Guillaume Janssens, Antoine Delor, John Aldo Lee, and Xavier Geets. Validation of the mid-position strat-

- egy for lung tumors in helical tomotherapy. *Radiotherapy and Oncology*, 110(3):529–537, 2014.
- [219] Miao Zhang, Wei Zou, and Boon-Keng Kevin Teo. Image guidance in proton therapy for lung cancer. *Translational lung cancer research*, 7(2):160, 2018.
- [220] Guillaume Landry and Chia-ho Hua. Current state and future applications of radiological image guidance for particle therapy. *Medical Physics*, 45(11):e1086–e1095, 2018.
- [221] Jihyun Yun, Eugene Yip, Zsolt Gabos, Keith Wachowicz, Satyapal Rathee, and BG Fallone. Neural-network based autocontouring algorithm for intrafractional lung-tumor tracking using linac-mr. *Medical physics*, 42(5):2296–2310, 2015.
- [222] Dan Nguyen, Xun Jia, David Sher, Mu-Han Lin, Zohaib Iqbal, Hui Liu, and Steve Jiang. 3d radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected u-net deep learning architecture. *Physics in medicine & Biology*, 64(6):065020, 2019.
- [223] Ana María Barragán-Montero, Dan Nguyen, Weiguo Lu, Mu-Han Lin, Roya Norouzi-Kandalan, Xavier Geets, Edmond Sterpin, and Steve Jiang. Three-dimensional dose prediction for lung imrt patients with deep neural networks: robust learning from heterogeneous beam configurations. *Medical physics*, 46(8):3679–3691, 2019.
- [224] Caroline Lavoie, Jane Higgins, Jean-Pierre Bissonnette, Lisa W Le, Alexander Sun, Anthony Brade, Andrew Hope, John Cho, and Andrea Bezjak. Volumetric image guidance using carina vs spine as registration landmarks for conventionally fractionated lung radiotherapy. *International Journal of Radiation Oncology* Biology* Physics*, 84(5):1086–1092, 2012.
- [225] Antje-Christin Knopf, Theodore S Hong, and Antony Lomax. Scanned proton radiotherapy for mobile targets—the effectiveness of re-scanning in the context of different treatment planning approaches and for different motion characteristics. *Physics in Medicine & Biology*, 56(22):7257, 2011.
- [226] Catarina Veiga, Guillaume Janssens, Ching-Ling Teng, Thomas Baudier, Lucian Hotoiu, Jamie R McClelland, Gary Royle, Liyong Lin, Lingshu Yin, James Metz, et al. First clinical investigation of cone beam computed tomography and deformable registration for adaptive proton therapy for lung cancer. *International Journal of Radiation Oncology* Biology* Physics*,

- 95(1):549–559, 2016.
- [227] Tara A Van De Water, Antony J Lomax, Hendrik P Bijl, Marije E De Jong, Cornelis Schilstra, Eugen B Hug, and Johannes A Langendijk. Potential benefits of scanned intensity-modulated proton therapy versus advanced photon therapy with regard to sparing of the salivary glands in oropharyngeal cancer. *International Journal of Radiation Oncology* Biology* Physics*, 79(4):1216–1224, 2011.
- [228] Charles B. Simone, David Ly, Tu D. Dan, John Ondos, Holly Ning, Arnaud Belard, John O’Connell, Robert W. Miller, and Nicole L. Simone. Comparison of intensity-modulated radiotherapy, adaptive radiotherapy, proton radiotherapy, and adaptive proton radiotherapy for treatment of locally advanced head and neck cancer. *Radiotherapy and Oncology*, 101(3):376–382, 2011.
- [229] Joanna Góra, Peter Kuess, Markus Stock, Piotr Andrzejewski, Barbara Knäusl, Brigita Paskeviciute, Gabriela Altorjai, and Dietmar Georg. ART for head and neck patients: On the difference between VMAT and IMPT. *Acta Oncol.*, 54(8):1166–1174, April 2015.
- [230] Tineke W.H. Meijer, Dan Scandurra, and Johannes A. Langendijk. Reduced radiation-induced toxicity by using proton therapy for the treatment of oropharyngeal cancer. *The British Journal of Radiology*, 93(1107):20190955, 2020. PMID: 31971818.
- [231] Hans Paul Van Der Laan, Tara A. Van De Water, Heleen E. Van Herpt, Miranda E.M.C. Christianen, Hendrik P. Bijl, Erik. W. Korevaar, Coen R. Rasch, Aart A. Van’t Veld, Arjen Van Der Schaaf, Cornelis Schilstra, and Johannes A. Langendijk. The potential of intensity-modulated proton radiotherapy to reduce swallowing dysfunction in the treatment of head and neck cancer: A planning comparative study. *Acta Oncologica*, 52(3):561 – 569, 2013. Cited by: 71.
- [232] Gohar S Manzar, Scott C Lester, David M Routman, William S Harmsen, Molly M Petersen, Jeff A Sloan, Daniel W Mundy, Ashley E Hunzeker, Adam C Amundson, Jeffrey L Anderson, Samir H Patel, Yolanda I Garces, Michele Y Halyard, Lisa A McGee, Michelle A Neben-Wittich, Daniel J Ma, Steven J Frank, Thomas J Whitaker, and Robert L Foote. Comparative analysis of acute toxicities and patient reported outcomes between intensity-modulated proton therapy (IMPT) and volumetric modulated arc therapy (VMAT) for the treatment of oropharyngeal cancer, 2020.

- [233] Y.C. Chou, S.P. Hung, C.E. Hsieh, Y.Y. Wu, and J.T.C. Chang. Intensity-modulated proton therapy reduces acute treatment-related toxicities for patients with nasopharyngeal cancer: A case-control propensity score match study with volumetric modulated arc therapy. *International Journal of Radiation Oncology*Biology*Physics*, 102(3, Supplement):e236–e237, 2018. Proceedings of the American Society for Radiation Oncology.
- [234] Macarena Cubillos-Mesías, Esther G C Troost, Fabian Lohaus, Linda Agolli, Maximilian Rehm, Christian Richter, and Kristin Stützer. Including anatomical variations in robust optimization for head and neck proton therapy can reduce the need of adaptation. *Radiother. Oncol.*, 131:127–134, February 2019.
- [235] Aafke C Kraan, Steven van de Water, David N Teguh, Abraham Al-Mamgani, Tom Madden, Hanne M Kooy, Ben J M Heijmen, and Mischa S Hoogeman. Dose uncertainties in IMPT for oropharyngeal cancer in the presence of anatomical, range, and setup errors. *Int. J. Radiat. Oncol. Biol. Phys.*, 87(5):888–896, December 2013.
- [236] Jia Yi Lim and Michelle Leech. Use of auto-segmentation in the delineation of target volumes and organs at risk in head and neck. *Acta Oncol.*, 55(7):799–806, July 2016.
- [237] Elena Borderías Villarroel, Xavier Geets, and Edmond Sterpin. Online adaptive dose restoration in intensity modulated proton therapy of lung cancer to account for inter-fractional density changes. *Phys Imaging Radiat Oncol*, 15:30–37, July 2020.
- [238] Thyrza Z Jagt, Sebastiaan Breedveld, Rens van Haveren, Ben J M Heijmen, and Mischa S Hoogeman. Online-adaptive versus robust IMPT for prostate cancer: How much can we gain? *Radiother. Oncol.*, 151:228–233, October 2020.
- [239] Gloria Vilches-Freixas, Mirko Unipan, Ilaria Rinaldi, Jonathan Martens, Erik Roijen, Isabel P Almeida, Esther Decabooter, and Geert Bosmans. Beam commissioning of the first compact proton therapy system with spot scanning and dynamic field collimation. *Br. J. Radiol.*, 93(1107):20190598, March 2020.
- [240] Emily Neubauer, Lei Dong, David S Followill, Adam S Garden, Laurence E Court, R Allen White, and Stephen F Kry. Assessment of shoulder position variation and its impact on IMRT and VMAT doses for head and neck

- cancer. *Radiat. Oncol.*, 7:19, February 2012.
- [241] Hidenobu Tachibana, Kana Motegi, and Shunsuke Moriya. Impact of shoulder deformation on volumetric modulated arc therapy doses for head and neck cancer, 2018.
- [242] M Kosmin, J Ledsam, B Romera-Paredes, R Mendes, S Moinuddin, D de Souza, L Gunn, C Kelly, C O Hughes, A Karthikesalingam, C Nutting, and R A Sharma. Rapid advances in auto-segmentation of organs at risk and target volumes in head and neck cancer. *Radiother. Oncol.*, 135:130–140, June 2019.
- [243] L Capelle, M Mackenzie, C Field, M Parliament, S Ghosh, and R Scrimger. Adaptive radiotherapy using helical tomotherapy for head and neck cancer in definitive and postoperative settings: initial results. *Clin. Oncol.*, 24(3):208–215, April 2012.
- [244] G. Guterres Marmitt, A. Pin, K. Ng Wei Siang, G. Janssens, K. Souris, M. Cohilis, J.A. Langendijk, S. Both, A. Knopf, and A. Meijers. Platform for automatic patient quality assurance via monte carlo simulations in proton therapy. *Physica Medica*, 70:49 – 57, 2020.
- [245] Kwangzoo Chung. A pilot study of the scanning beam quality assurance using machine log files in proton beam therapy. *Progress in Medical Physics*, 28:129, 01 2017.
- [246] Carlos E Cardenas, Beth M Beadle, Adam S Garden, Heath D Skinner, Jinzhong Yang, Dong Joo Rhee, Rachel E McCarroll, Tucker J Netherton, Skylar S Gay, Lifei Zhang, and Laurence E Court. Generating High-Quality lymph node clinical target volumes for head and neck cancer radiation therapy using a fully automated deep Learning-Based approach. *Int. J. Radiat. Oncol. Biol. Phys.*, 109(3):801–812, March 2021.
- [247] Jordan Wong, Allan Fong, Nevin McVicar, Sally Smith, Joshua Giambattista, Derek Wells, Carter Kolbeck, Jonathan Giambattista, Lovedeep Gondara, and Abraham Alexander. Comparing deep learning-based auto-segmentation of organs at risk and clinical target volumes to expert inter-observer variability in radiotherapy planning. *Radiother. Oncol.*, 144:152–158, March 2020.
- [248] F Guerreiro, E Seravalli, G O Janssens, J H Maduro, A C Knopf, J A Langendijk, B W Raaymakers, and C Kontaxis. Deep learning prediction

- of proton and photon dose distributions for paediatric abdominal tumours. *Radiother. Oncol.*, 156:36–42, March 2021.
- [249] Kristoffer Petersson, Per Nilsson, Per Engström, Tommy Knöös, and Crister Ceberg. Evaluation of dual-arc VMAT radiotherapy treatment plans automatically generated via dose mimicking. *Acta Oncol.*, 55(4):523–525, 2016.
- [250] P Meyer, M-C Biston, C Khamphan, T Marghani, J Mazurier, V Bodez, L Fezzani, P A Rigaud, G Sidorski, L Simon, and C Robert. Automation in radiotherapy treatment planning: Examples of use in clinical practice and future trends for a complete automated workflow. *Cancer Radiother.*, 25(6-7):617–622, October 2021.
- [251] Yaorong Ge and Q Jackie Wu. Knowledge-based planning for intensity-modulated radiation therapy: a review of data-driven approaches. *Medical physics*, 46(6):2760–2775, 2019.
- [252] Aaron Babier, Rafid Mahmood, Andrea L McNiven, Adam Diamant, and Timothy CY Chan. The importance of evaluating the complete automated knowledge-based planning pipeline. *Physica Medica*, 72:73–79, 2020.
- [253] Ana M Barragán-Montero, Melissa Thomas, Gilles Defraene, Steven Michiels, Karin Haustermans, John A Lee, and Edmond Sterpin. Deep learning dose prediction for imrt of esophageal cancer: The effect of data quality and quantity on model performance. *Physica Medica*, 83:52–63, 2021.
- [254] Jiawei Fan, Jiazhou Wang, Zhi Chen, Chaosu Hu, Zhen Zhang, and Weigang Hu. Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique. *Medical physics*, 46(1):370–381, 2019.
- [255] Camille Draguet, Ana M Barragán-Montero, Macarena Chocan Vera, Melissa Thomas, Pieter Populaire, Gilles Defraene, Karin Haustermans, John A Lee, and Edmond Sterpin. Automated clinical decision support system with deep learning dose prediction and ntcp models to evaluate treatment complications in patients with esophageal cancer. *Radiotherapy and Oncology*, 2022.
- [256] Wei Wang, Yu Chang, Yilin Liu, Zhikai Liang, Yicheng Liao, Bin Qin, Xu Liu, and Zhiyong Yang. Feasibility study of fast intensity-modulated proton therapy dose prediction method using deep neural networks for prostate

- cancer. *Medical Physics*, 2022.
- [257] Martin Stuschke, Andreas Kaiser, Jehad Abu Jawad, Christoph Pöttgen, Sabine Levegrün, and Jonathan Farr. Multi-scenario based robust intensity-modulated proton therapy (impt) plans can account for set-up errors more effectively in terms of normal tissue sparing than planning target volume (ptv) based intensity-modulated photon plans in the head and neck region. *Radiation Oncology*, 8(1):1–5, 2013.
- [258] Oskar Eriksson and Tianfang Zhang. Robust automated radiation therapy treatment planning using scenario-specific dose prediction and robust dose mimicking. *Medical Physics*, 2022.
- [259] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.
- [260] RaySearch Laboratories. Machine learning planning model RSL-IMPT-OROPHARYNX-7000-SIB (3.0)*. <https://www.raysearchlabs.com/496c6e/siteassets/raystation-landing-page/machine-learning/available-pretrained-models/rsl-impt-oropharynx-7000-sib-3.0-model-flyer.pdf>. Accessed: 2022-7-14.
- [261] Elena Borderías-Villarroel, Vicki Taasti, Wouter Van Elmpt, S Teruel-Rivas, X Geets, and E Sterpin. Evaluation of the clinical value of automatic on-line dose restoration for adaptive proton therapy of head and neck cancer. *Radiotherapy and Oncology*, 170:190–197, 2022.
- [262] Michael Goitein and Andrzej Niemierko. Intensity modulated therapy and inhomogeneous dose to the tumor: a note of caution. *International journal of radiation oncology, biology, physics*, 36(2):519–522, 1996.
- [263] Roel GJ Kierkels, Albin Fredriksson, Stefan Both, Johannes A Langendijk, Daniel Scandurra, and Erik W Korevaar. Automated robust proton planning using dose-volume histogram-based mimicking of the photon reference dose and reducing organ at risk dose optimization. *International Journal of Radiation Oncology* Biology* Physics*, 103(1):251–258, 2019.
- [264] Junjie Hu, Ying Song, Qiang Wang, Sen Bai, and Zhang Yi. Incorporating historical sub-optimal deep neural networks for dose prediction in radiother-

- apy. *Medical Image Analysis*, 67:101886, 2021.
- [265] Liesbeth Vandewinckele, Siri Willems, Maarten Lambrecht, Patrick Berkovic, Frederik Maes, and Wouter Crijns. Treatment plan prediction for lung imrt using deep learning based fluence map generation. *Physica Medica*, 99:44–54, 2022.
- [266] Lin Ma, Mingli Chen, Xuejun Gu, and Weiguo Lu. Deep learning-based inverse mapping for fluence map prediction. *Physics in Medicine & Biology*, 65(23):235035, 2020.
- [267] Wentao Wang, Yang Sheng, Manisha Palta, Brian Czito, Christopher Willett, Martin Hito, Fang-Fang Yin, Qiuwen Wu, Yaorong Ge, and Q Jackie Wu. Deep learning-based fluence map prediction for pancreas stereotactic body radiation therapy with simultaneous integrated boost. *Advances in radiation oncology*, 6(4):100672, 2021.
- [268] Huet-Dastarac M, Barragan-Montero A. M., Michiels S., Teruel-Rivas S., Ozan H, Sterpin E, Lee J. Automatic decision tool for head and neck patient referral based on dose prediction with deep learning.
- [269] David J Sher, Andrew Godley, Yang Park, Colin Carpenter, Marc Nash, Hasti Hesami, Xinran Zhong, and Mu-Han Lin. Prospective study of artificial intelligence-based decision support to improve head and neck radiotherapy plan quality. *Clinical and translational radiation oncology*, 29:65–70, 2021.
- [270] Christopher Kurz, Matteo Maspero, Mark H F Savenije, Guillaume Landry, Florian Kamp, Marco Pinto, Minglun Li, Katia Parodi, Claus Belka, and Cornelis A T van den Berg. CBCT correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation. *Phys. Med. Biol.*, 64(22):225004, November 2019.
- [271] Liyuan Chen, Xiao Liang, Chenyang Shen, Dan Nguyen, Steve Jiang, and Jing Wang. Synthetic ct generation from cbct images via unsupervised deep learning. *Physics in Medicine & Biology*, 66(11):115019, 2021.
- [272] Steven Van De Water, Iris van Dam, Dennis R Schaart, Abraham Al-Mamgani, Ben JM Heijmen, and Mischa S Hoogeman. The price of robustness; impact of worst-case optimization on organ-at-risk dose and complication probability in intensity-modulated proton therapy for oropharyngeal cancer patients. *Radiotherapy and Oncology*, 120(1):56–62, 2016.

- [273] Stephen J Gardner, Weihua Mao, Chang Liu, Ibrahim Aref, Mohamed Elshaikh, Joon K Lee, Deepak Pradhan, Benjamin Movsas, Indrin J Chetty, and Farzan Siddiqui. Improvements in cbct image quality using a novel iterative reconstruction algorithm: a clinical evaluation. *Advances in radiation oncology*, 4(2):390–400, 2019.
- [274] Talia Jarema and Trent Aland. Using the iterative kv cbct reconstruction on the varian halcyon linear accelerator for radiation therapy planning for pelvis patients. *Physica Medica*, 68:112–116, 2019.
- [275] Valentina Giacometti, Alan R Hounsell, and Conor K McGarry. A review of dose calculation approaches with cone beam ct in photon and proton therapy. *Physica Medica*, 76:243–276, 2020.
- [276] Ola Weistrand and Stina Svensson. The anaconda algorithm for deformable image registration in radiotherapy. *Medical physics*, 42(1):40–53, 2015.
- [277] Johannes A Langendijk, Frank JP Hoebbers, Martin A De Jong, Patricia Doornaert, Chris HJ Terhaard, Roel JHM Steenbakkers, Olga Hamming-Vrieze, Jeroen B Van De Kamer, Wilko FAR Verbakel, Fatma Keskin-Cambay, et al. National protocol for model-based selection for proton therapy in head and neck cancer. *International journal of particle therapy*, 8(1):354–365, 2021.
- [278] Sara Thörnqvist, Liv B Hysing, Laura Tuomikoski, Anne Vestergaard, Kari Tanderup, Ludvig P Muren, and Ben J M Heijmen. Adaptive radiotherapy strategies for pelvic tumors – a systematic review of clinical implementations, 2016.
- [279] Eliott Brion, Jean Léger, Umair Javaid, John Lee, Christophe De Vleeschouwer, and Benoît Macq. Using planning cts to enhance cnn-based bladder segmentation on cone beam ct. In *Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 10951, pages 410–417. SPIE, 2019.
- [280] Eliott Brion, Jean Léger, Ana M Barragán-Montero, Nicolas Meert, John A Lee, and Benoît Macq. Domain adversarial networks and intensity-based data augmentation for male pelvic organ segmentation in cone beam ct. *Computers in Biology and Medicine*, 131:104269, 2021.
- [281] Heng Li, William T Hrinivich, Hao Chen, Khadija Sheikh, Meng Wei Ho, Rachel Ger, Dezhi Liu, Russell Kenneth Hales, Khanh Ranh Voong, Aditya

- Halhore, et al. Evaluating proton dose and associated range uncertainty using daily cone-beam ct. *Frontiers in Oncology*, 12, 2022.
- [282] Lian Zhang, Zhi Wang, Chengyu Shi, Tengfei Long, and X George Xu. The impact of robustness of deformable image registration on contour propagation and dose accumulation for head and neck adaptive radiotherapy. *Journal of Applied Clinical Medical Physics*, 19(4):185–194, 2018.
- [283] Catarina Veiga, Ana Mónica Lourenço, Syed Mouinuddin, Marcel Van Herk, Marc Modat, Sébastien Ourselin, Gary Royle, and Jamie R McClelland. Toward adaptive radiotherapy for head and neck patients: uncertainties in dose warping due to the choice of deformable registration algorithm. *Medical Physics*, 42(2):760–769, 2015.
- [284] Yong Gan, Johannes A Langendijk, Edwin Oldehinkel, Daniel Scandurra, Nanna M Sijtsema, Zhixiong Lin, Stefan Both, and Charlotte L Brouwer. A novel semi auto-segmentation method for accurate dose and ntcp evaluation in adaptive head and neck radiotherapy. *Radiotherapy and Oncology*, 164:167–174, 2021.
- [285] William J Beasley, Alan McWilliam, Nicholas J Slevin, Ranald I Mackay, and Marcel Van Herk. An automated workflow for patient-specific quality control of contour propagation. *Physics in Medicine & Biology*, 61(24):8577, 2016.
- [286] G Guterres Marmitt, A Pin, K Ng Wei Siang, G Janssens, Kevin Souris, Marie Cohilis, JA Langendijk, S Both, A Knopf, and A Meijers. Platform for automatic patient quality assurance via monte carlo simulations in proton therapy. *Physica Medica*, 70:49–57, 2020.
- [287] Vicki Trier Taasti, Colien Hazelaar, Femke Vaassen, Ana Vaniqui, Karolien Verhoeven, Frank Hoebbers, Wouter van Elmpt, Richard Canters, and Mirko Unipan. Clinical implementation and validation of an automated adaptive workflow for proton therapy. *Physics and imaging in radiation oncology*, 2022.
- [288] Lena Nenoff, Michael Matter, Marjolaine Charmillot, Serge Krier, Klara Uher, Damien Charles Weber, Antony John Lomax, and Francesca Albertini. Experimental validation of daily adaptive proton therapy. *Physics in Medicine & Biology*, 66(20):205010, 2021.
- [289] Steven Van De Water, Iris van Dam, Dennis R Schaart, Abraham Al-Mamgani, Ben JM Heijmen, and Mischa S Hoogeman. The price of ro-

- bustness; impact of worst-case optimization on organ-at-risk dose and complication probability in intensity-modulated proton therapy for oropharyngeal cancer patients. *Radiotherapy and Oncology*, 120(1):56–62, 2016.
- [290] Dirk Wagenaar, Roel GJ Kierkels, Arjen van der Schaaf, Arturs Meijers, Daniel Scandurra, Nanna M Sijtsema, Erik W Korevaar, Roel JHM Steenbakkers, Antje C Knopf, Johannes A Langendijk, et al. Head and neck impt probabilistic dose accumulation: Feasibility of a 2 mm setup uncertainty setting. *Radiotherapy and Oncology*, 154:45–52, 2021.
- [291] Nils Peters, Patrick Wohlfahrt, Christian Hofmann, Christian Möhler, Stefan Menkel, Maria Tschiche, Mechthild Krause, Esther GC Troost, Wolfgang Enghardt, and Christian Richter. Reduction of clinical safety margins in proton therapy enabled by the clinical implementation of dual-energy ct for direct stopping-power prediction. *Radiotherapy and Oncology*, 166:71–78, 2022.
- [292] M Dohopolski, B Choi, B Meng, J Visak, X Zhong, JS Kim, E Inam, V Avkshtol, DH Moon, DJ Sher, et al. Dosimetric impact of simulated daily adaptive radiotherapy with significantly reduced setup margins in the definitive treatment of head and neck cancer. *International Journal of Radiation Oncology, Biology, Physics*, 114(3):e590, 2022.
- [293] Lina M Åström, Claus P Behrens, Katrine Smedegaard Storm, Patrik Sibolt, and Eva Serup-Hansen. Online adaptive radiotherapy of anal cancer: Normal tissue sparing, target propagation methods, and first clinical experience. *Radiotherapy and Oncology*, 176:92–98, 2022.
- [294] Simon van Kranen, Olga Hamming-Vrieze, Annelisa Wolf, Eugène Damen, Marcel van Herk, and Jan-Jakob Sonke. Head and neck margin reduction with adaptive radiation therapy: robustness of treatment plans against anatomy changes. *International Journal of Radiation Oncology* Biology* Physics*, 96(3):653–660, 2016.
- [295] Joe Y Chang, Lei Dong, Helen Liu, George Starkschall, Peter Balter, Radhe Mohan, Zhongxing Liao, James D Cox, and Ritsuko Komaki. Image-guided radiation therapy for non-small cell lung cancer. *Journal of Thoracic Oncology*, 3(2):177–186, 2008.
- [296] Olga Hamming-Vrieze, Simon R van Kranen, Wilma D Heemsbergen, Charlotte AH Lange, Michiel WM van den Brekel, Marcel Verheij, Coen RN Rasch, and Jan Jakob Sonke. Analysis of gtv reduction during radiotherapy

- for oropharyngeal cancer: Implications for adaptive radiotherapy. *Radiotherapy and Oncology*, 122(2):224–228, 2017.
- [297] Hamid Raziee, Fabio Y Moraes, Jure Murgic, Melvin LK Chua, Melania Pintilie, Peter Chung, Cynthia Ménard, Andrew Bayley, Mary Gospodarowicz, Padraig Warde, et al. Improved outcomes with dose escalation in localized prostate cancer treated with precision image-guided radiotherapy. *Radiotherapy and Oncology*, 123(3):459–465, 2017.
- [298] Mitchell Machtay, Kyoungwha Bae, Benjamin Movsas, Rebecca Paulus, Elizabeth M Gore, Ritsuko Komaki, Kathy Albain, William T Sause, and Walter J Curran. Higher biologically effective dose of radiotherapy is associated with improved outcomes for locally advanced non-small cell lung carcinoma treated with chemoradiation: an analysis of the radiation therapy oncology group. *International Journal of Radiation Oncology* Biology* Physics*, 82(1):425–434, 2012.
- [299] Mark R Storey, Alan Pollack, Gunar Zagars, Lewis Smith, John Antolak, and Isaac Rosen. Complications from radiotherapy dose escalation in prostate cancer: preliminary results of a randomized trial. *International Journal of Radiation Oncology* Biology* Physics*, 48(3):635–642, 2000.
- [300] Jeffrey D Bradley, Rebecca Paulus, Ritsuko Komaki, Gregory Masters, George Blumenschein, Steven Schild, Jeffrey Bogart, Chen Hu, Kenneth Forster, Anthony Magliocco, et al. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage iiia or iiib non-small-cell lung cancer (rtog 0617): a randomised, two-by-two factorial phase 3 study. *The lancet oncology*, 16(2):187–199, 2015.
- [301] Christina K Speirs, Todd A DeWees, Sana Rehman, Alerson Molotievschi, Maria A Velez, Daniel Mullen, Sandra Fergus, Marco Trovo, Jeffrey D Bradley, and Cliff G Robinson. Heart dose is an independent dosimetric predictor of overall survival in locally advanced non-small cell lung cancer. *Journal of Thoracic Oncology*, 12(2):293–301, 2017.
- [302] Lisanne V van Dijk, Steven J Frank, Ying Yuan, Brandon Gunn, Amy C Moreno, Abdallah SR Mohamed, Kathryn E Preston, Yun Qing, Michael T Spiotto, William H Morrison, et al. Proton image-guided radiation assignment for therapeutic escalation via selection of locally advanced head and neck cancer patients [pirates]: A phase i safety and feasibility trial of mri-guided adaptive particle radiotherapy. *Clinical and translational radiation*

- oncology*, 32:35–40, 2022.
- [303] Laura Grocutt, Claire Paterson, and Ronan M Valentine. Adaptive dose escalated radiotherapy in oropharyngeal cancers: a treatment planning feasibility study. *Radiation Oncology*, 17(1):1–11, 2022.
- [304] Joey P Cheung, Peter C Park, Laurence E Court, X Ronald Zhu, Rajat J Kudchadker, Steven J Frank, and Lei Dong. A novel dose-based positioning method for ct image-guided proton therapy. *Medical physics*, 40(5):051714, 2013.
- [305] Lena Nenoff, Cássia O Ribeiro, Michael Matter, Luana Hafner, Mirjana Josipovic, Johannes A Langendijk, Gitte F Persson, Marc Walser, Damien Charles Weber, Antony John Lomax, et al. Deformable image registration uncertainty for inter-fractional dose accumulation of lung cancer proton therapy. *Radiotherapy and Oncology*, 147:178–185, 2020.
- [306] Ditte Sloth Møller, Marianne Ingerslev Holt, Markus Alber, Marie Tvillum, Azza Ahmed Khalil, Marianne Marquard Knap, and Lone Hoffmann. Adaptive radiotherapy for advanced lung cancer ensures target coverage and decreases lung dose. *Radiother. Oncol.*, 121(1):32–38, October 2016.
- [307] Muhammad Shamshad, Ditte Sloth Møller, Hanna Rahbek Mortensen, Mai Lykkegaard Ehmsen, Maria Fuglsang Jensen, and Lone Hoffmann. Bone versus soft-tissue setup in proton therapy for patients with oesophageal cancer. *Acta Oncologica*, 61(8):994–1003, 2022.
- [308] G Brandon Gunn, Pierre Blanchard, Adam S Garden, X Ronald Zhu, C David Fuller, Abdallah S Mohamed, William H Morrison, Jack Phan, Beth M Beadle, Heath D Skinner, et al. Clinical outcomes and patterns of disease recurrence after intensity modulated proton therapy for oropharyngeal squamous carcinoma. *International Journal of Radiation Oncology* Biology* Physics*, 95(1):360–367, 2016.
- [309] L Nenoff, A Köthe, M Matter, E Amaya, M Josipovic, Antje Knopf, G Persson, Cássia O Ribeiro, S Safai, Sabine Visser, et al. Tcp and ntcp calculations based on treatment doses instead of planned doses for daily adaptive proton therapy of lung cancer. *International journal of radiation oncology, biology, physics*, 111(3):e133, 2021.
- [310] Djoya Hattu, Jolein Mannens, Michel Öllers, Judith van Loon, Dirk De Ruysscher, and Wouter van Elmpt. A traffic light protocol workflow

- for image-guided adaptive radiotherapy in lung cancer patients. *Radiotherapy and Oncology*, 175:152–158, 2022.
- [311] Monica Buijs, Floris Pos, Marloes Frantzen-Steneker, Maddalena Rossi, Peter Remeijer, and Folkert Koetsveld. Take action protocol: a radiation therapist led approach to act on anatomical changes seen on cbct. *Technical innovations & patient support in radiation oncology*, 17:71–77, 2021.
- [312] Nicolas F Chaves-de Plaza, Prerak Mody, Klaus Hildebrandt, Marius Starling, Huib de Ridder, and Rene van Egmond. Towards fast ai-infused human-centered contouring workflows for adaptive proton therapy in the head and neck. *arXiv preprint arXiv:2208.04675*, 2022.
- [313] HA McNair, E Joyce, G O’Gara, M Jackson, B Peet, RA Huddart, and T Wiseman. Radiographer-led online image guided adaptive radiotherapy: A qualitative investigation of the therapeutic radiographer role. *Radiography*, 27(4):1085–1093, 2021.
- [314] Meegan Shepherd, Siobhan Graham, Amy Ward, Lissane Zwart, Bin Cai, Charlotte Shelley, and Jeremy Booth. Pathway for radiation therapists online advanced adapter training and credentialing. *Technical Innovations & Patient Support in Radiation Oncology*, 20:54–60, 2021.
- [315] D Daal, L Haverkate, L ten Asbroek-Zwolsman, L Zwart, E van Dieren, and E de Wit. Pd-0940 cbct-guided online adaptive radiotherapy: implementation of an rtt-led workflow. *Radiotherapy and Oncology*, 161:S783, 2021.
- [316] Thomas Willigenburg, Daan M de Muinck Keizer, Max Peters, An Claes, Jan JW Lagendijk, Hans CJ de Boer, et al. Evaluation of daily online contour adaptation by radiation therapists for prostate cancer treatment on an mri-guided linear accelerator. *Clinical and translational radiation oncology*, 27:50–56, 2021.
- [317] Marnix JA Rasing, Gonda G Sikkes, Nicole GPM Vissers, Alexis NTJ Kotte, Joske H Boudewijn, Patricia AH Doornaert, Wietse SC Eppinga, Martijn Intven, Reijer HA Rutgers, Annick Scheeren, et al. Online adaptive mr-guided radiotherapy: Conformity of contour adaptation for prostate cancer, rectal cancer and lymph node oligometastases among radiation therapists and radiation oncologists. *Technical innovations & patient support in radiation oncology*, 23:33–40, 2022.

- [318] Lena Nenoff, Michael Matter, Enrique Javier Amaya, Mirjana Josipovic, Antje-Christin Knopf, Antony John Lomax, Gitte F Persson, Cássia O Ribeiro, Sabine Visser, Marc Walser, et al. Dosimetric influence of deformable image registration uncertainties on propagated structures for on-line daily adaptive proton therapy of lung cancer patients. *Radiotherapy and oncology*, 159:136–143, 2021.
- [319] Thomas Berger, Jérémy Godart, Thyrza Jagt, Anders Schwartz Vittrup, Lars Ulrik Fokdal, Jacob Christian Lindegaard, Nina Boje Kibsgaard Jensen, Andras Zolnay, Dominique Reijtenbagh, Petra Trnkova, et al. Dosimetric impact of intrafraction motion in online-adaptive intensity modulated proton therapy for cervical cancer. *International Journal of Radiation Oncology* Biology* Physics*, 109(5):1580–1587, 2021.
- [320] Bradley M Oborn, Stephen Dowdell, Peter E Metcalfe, Stuart Crozier, Radhe Mohan, and Paul J Keall. Future of medical physics: real-time mri-guided proton therapy. *Medical physics*, 44(8):e77–e90, 2017.
- [321] Maryam Moteabbed, Julien Smeets, Theodore S Hong, Guillaume Janssens, Rudi Labarbe, John A Wolfgang, and Thomas R Bortfeld. Toward mr-integrated proton therapy: modeling the potential benefits for liver tumors. *Physics in Medicine & Biology*, 66(19):195004, 2021.
- [322] Steven Van De Water, Sairos Safai, Jacobus M Schippers, Damien C Weber, and Antony J Lomax. Towards flash proton therapy: the impact of treatment planning and machine characteristics on achievable dose rates. *Acta oncologica*, 58(10):1463–1469, 2019.
- [323] Manuela Buonanno, Veljko Grilj, and David J Brenner. Biological effects in normal cells exposed to flash dose rate protons. *Radiotherapy and Oncology*, 139:51–55, 2019.
- [324] Pierre Lansonneur, Hamid Mammar, Catherine Nauraye, Annalisa Patriarca, Eric Hierso, Rémi Dendale, Yolanda Prezado, and Ludovic De Marzi. First proton minibeam radiation therapy treatment plan evaluation. *Scientific reports*, 10(1):1–8, 2020.
- [325] Thomas Henry, Ana Ureba, Alexander Valdman, and Albert Siegbahn. Proton grid therapy: a proof-of-concept study. *Technology in Cancer Research & Treatment*, 16(6):749–757, 2017.
- [326] Xiaoqiang Li, Gang Liu, Guillaume Janssens, Olivier De Wilde, Vincent

- Bossier, Xavier Lerot, Antoine Pouppez, Di Yan, Craig Stevens, Peyman Kabolizadeh, et al. The first prototype of spot-scanning proton arc treatment delivery. *Radiotherapy and Oncology*, 137:130–136, 2019.
- [327] Priya Bhatnagar, Manil Subesinghe, Chirag Patel, Robin Prestwich, and Andrew F Scarsbrook. Functional imaging for radiation treatment planning, response assessment, and adaptive therapy in head and neck cancer. *Radiographics*, 33(7):1909–1929, 2013.
- [328] Narek Shaverdian, Yingli Yang, Peng Hu, Steven Hart, Ke Sheng, James Lamb, Minsong Cao, Nzhde Agazaryan, David Thomas, Michael Steinberg, et al. Feasibility evaluation of diffusion-weighted imaging using an integrated mri-radiotherapy system for response assessment to neoadjuvant therapy in rectal cancer. *The British Journal of Radiology*, 90(1071):20160739, 2017.