Adaptive Fractionation for Reducing Number of Fractions

Master Thesis - Medical Physics
University of Zurich, Department of Physics

Author
Janic Tom Weber

Date
March 02, 2023

Research Group Leader
Prof. Dr. Jan Unkelbach

Supervisors
Dr. Roman Ludwig
Dr. Riccardo Dal Bello

Abstract

Objective: Fractionated radiotherapy typically delivers the same dose in each fraction (Uniform Fractionation). Adaptive Fractionation is a technique proposed by Pérez Haas [1] to exploit inter-fractional motion by increasing the dose on days when the distance of tumor and dose-limiting OAR is large and decreasing the dose on days when the distance is small. For favourable patient geometries where distances are large, Adaptive Fractionation has shown to deliver small residual doses in final or close-to-final fractions. Developed is an extension of the Adaptive Fractionation model to minimise number of fractions used for a treatment, to prevent applying such small residual doses in the final or close-to-final fractions for favourable patient geometries and use the vacant treatment allocation slot for additional patients: On favourable days the dose is further increased to possibly finish the treatment in earlier fractions and on unfavourable days dose modification is conformed to standard Adaptive Fractionation utilising the prescribed maximum number of fractions. The extended concept is evaluated for patients with pancreas, adrenal glands and prostate tumors previously treated at the MR-Linac in 5 fractions with ablative dose.

Approach: Given daily adapted treatment plans, inter-fractional changes are quantified by sparing factors δ_t defined as the OAR-to-tumor dose ratio. The key problem of Adaptive Fractionation is to decide on the dose to deliver in fraction t, given δ_t and the dose delivered in previous fractions, but not knowing future δ_ts. Optimal doses that minimise the expected biologically effective dose in the OAR BED_3 and the number of fractions, while delivering a minimal BED_{10} tumor dose prescription, are computed using dynamic programming. Assumed is a normal distribution over δ with mean and variance estimated from previously observed patient-specific δ_ts for modelling sparing factor distribution. Collected were data of 30 patients from the MR-Linac treatment planning system for pancreas, adrenal glands and prostate cancer. The algorithm is evaluated retrospectively for two patients with pancreas tumor and one patient with tumor in the adrenal glands.

Main Results: In two patients with pancreatic cancer reducing number of fractions with Adaptive Fractionation resulted in a BED_3 decrease of 16.7 Gy respectively 2.8 Gy compared to Uniform Fractionation. The treatment reduced the number of fractions to 2 respectively to 4 number of fractions. In one patient with cancer in adrenal glands, reducing number of fractions with Adaptive Fractionation led to no reduction in number of fractions and a BED_3 increase of 2.8 Gy compared to Uniform Fractionation, due to an unfavourable planning sparing factor, that is used for estimating the mean of δ.

1 Introduction

Radiotherapy aims to use ionising radiation to damage and destroy cancerous tissue. The MR-Linac generates photons with an energy of up to 6 MeV by accelerating electrons onto a target, where they collide and are decelerated. This process results in the production of a beam of bremsstrahlung from the high kinetic energy of the electrons [2]. During their interaction with matter, photons ionise molecules within cells. The ionised electrons are responsible for most of the biological damage caused by radiation. This is because these electrons can cause further ionisations in the molecules they collide with, as they move through the tissue [3].
Ionising radiation can induce lethal cell damage by forming highly reactive radicals inside the cell nucleus that chemically break bonds in molecules. Damages may cause loss of reproducibility or even inflict apoptosis in targeted cells. In general both malignant and benign cells are exposed to radiation in radiotherapy treatment and thus both experience former effects. Much of the motivation for improving the technology of radiation therapy stems from the desire to increase normal tissue complication sparing. This means maintaining functional integrity of irradiated normal tissue and reducing prevalence of normal tissue complication occurrence. There are two elements to the strategy of normal tissue sparing. First element is the existence of a difference in the radiation response of benign and malignant cells. In treatments this difference allows preservation of functional integrity in normal tissue included in the target volume. Difference in this response between malignant and benign cells is assumed to be due to repair kinetics and cell repopulation. To exploit this differential effect the dose is fractionated, that is delivered in small daily increments. Second element of the strategy which is not further discussed here involves the reduction of the dose delivered to normal tissues that are spatially separated from the tumor. [4]
Fractionated therapy is an important rationale for local therapy. Employing fractionation, normal tissue can tolerate higher doses allowing the therapeutic ratio to be increased. One problem of fractionated therapy is that daily treatment over an extended period can result in changes to the geometry of non-stationary organs such as the intestines. Motion of tumors and organs at risk (OAR) in between fractions greatly impairs dose conformity and is generally assumed to degrade the quality of treatments. To account for impaired conformity caused by motion, the target volume can be extended with a safety margin. Safety margins in turn worsen the trade-off between tumor coverage and normal tissue sparing that would be feasible without motion [5]. To understand the target volume definition and their extensions one has to understand how treatment plans for irradiation are created.

2 Theory

2.1 Treatment Planning

To develop an irradiation plan for radiotherapy, it is first necessary to identify and segment the tumor from normal tissue. This is typically done using computed tomography (CT) scans, which provide detailed anatomical information and electron density information that can be used to calculate dose. In some cases, additional imaging modalities such as magnetic resonance (MR) and positron emission tomography (PET) may be used to improve the accuracy of tumor localisation and differentiation from normal tissue with similar density. Once the tumor has been identified and outlined, it is expanded to create specific target volumes that are nested within one another. [3]
The gross tumor volume (GTV) represents the primary tumor mass as seen on clinical examination or imaging. Extending the GTV results in the clinical target volume (CTV), which includes any additional microscopic disease. The final extension is the planning target volume (PTV), which accounts for potential errors in positioning during treatment or changes in the size or shape of the tumor or organs. In case no additional CT or MR scans are taken during the course of treatment (which can span several weeks), the PTV margin is essential to account for any uncertainties that may arise due to geometric variations. In addition to outlining the target volumes, it is also necessary to identify and outline any organs at risk (OARs) during treatment planning. OARs are defined as "those normal tissues which lie adjacent to tumors and may therefore be included within treated volumes, with a risk that the radiation may impair their normal functioning" [3]. These organs may overlap with the PTV, which can significantly constrain the tumor prescription dose in order to avoid damaging the OAR. There may also be multiple OARs in proximity to the target volume, each of which imposes its own dose constraints.
Once the tumor and OARs have been outlined, an optimal dose distribution can be calculated. This process involves setting constraints and objectives based on the prescribed tumor dose and OAR dose limits. Under a given set of beam angles and interaction model, an optimisation algorithm is used to compute the optimal achievable dose distribution. [6]

2.2 Image guided radiotherapy

Central to the advances in radiotherapy delivery is the development of medical imaging technologies that provide the 3D delineation of the target volumes and OAR. The development of state-of-the-art imaging methods have enabled modern radiotherapy to develop into a highly personalised, tailored treatment [7]. Technologies for image guided radiotherapy include CT, MRI, and PET imaging, where MR images offers superior soft-tissue contrast compared to CT [8]. Image functionalities provide additional information, that can be used to confirm patient positioning, mon-

itor and mitigate anatomical changes. This enables conforming the treatment plan to new patient geometry [7]. However, the conformal dose distribution is only achieved on the reference image which is the planning image, a snapshot of patient anatomy. Inter-fractional and intra-factional anatomy changes from the planning image may deteriorate the dose conformity in actual delivered dose distribution [9]. The use of more frequent on-line imaging, such as cone beam computerised tomography [10] and integrated magnetic resonance imaging [11], allows to detect these anatomic changes and aid in correcting or minimising the effect of changes in geometry. Image guidance is used to define the target more accurately and helps minimising the margins of the PTV which give rise to stereotactic body radiotherapy [12]. In each fraction through the course of the SBRT treatment at our MR-Linac, the patient is imaged with MRI. The MR image is aligned with deformable image registration to the reference planning CT image. In the presence of large geometry changes compared to reference it is decided whether the dose plan needs to be reoptimised to ensure dose conformity.

2.3 Uniform Fractionation

As mentioned before one strategy of normal tissue sparing involves fractionation. In uniform fractionation the prescribed dose is divided into dose fractions of equal size, which are delivered in multiple sessions. To enable comparison between different fractionation schedules the biological effective dose (BED) model is used. More precisely, BED allows determining iso-effective dose fractionation schedules. It is regarded as a measure of the biological dose delivered by a particular combination of equal dose per fraction D/n and total dose D in n fractions to a given tissue [13]. It is defined as

$$B = D\left(1 + \frac{D/n}{\alpha/\beta}\right)$$

where α and β are coefficient that stem from the Track-Event-Model which in low dose regions in approximated with the BED model. The α/β ratio is a measure the repair capability : cells with a high α/β values respond earlier i.e. show reactions such as proliferative impairment or loss of function earlier. Thus, two adjacent tissues with different α/β ratio values, each receiving the same dose and fractionation, will be associated with different BEDs. This does not necessarily mean that one tissue sustains more biological damage than the other. [13]

3 Prior Work

Fractionated therapy typically delivers equal dose in every fraction to the tumor when treated with uniform fractionation. Since the plan is fractionated and treatment spans several days or weeks, inter-fractional motion arises. This motion is seen as a handicap to treatments as it can degrade the quality of dose plans and calls for mitigation strategies such as adaptive radiotherapy. Rather than viewing inter-fractional motion as a handicap, it can be seen as an opportunity: the variation in geometry due to motion may be exploited to achieve better treatment quality compared to uniform fractionation.

3.1 Concept

Adaptive Fractionation [14][15][16] is one approach to exploit inter-fractional motion. In a recently published work from Pérez Haas et al. [1] an approach to adaptive fractionation has been devised and demonstrated with patient data from an MR-Linac. Adaptive Fractionation extends the existing adaptive therapy paradigm to modification of the tumor dose in each fraction: the dose is increased on favourable treatment days, i.e. when the distance between tumor and dose-limiting OAR is relatively large; and the dose is reduced for unfavourable geometries, i.e. when the tumor and OAR are closer. Thereby, the ratio between total dose delivered to the OAR versus total dose delivered to the tumor may be improved compared to uniform treatments that deliver the same dose in each fraction [1]. The difficulty comes from not knowing whether the remaining future fractions will have favourable or unfavourable patient geometries. The sparing factors are random variables with an estimated probability distribution but the exact future values are unknown. In the following the preliminary work is explained to get an understanding of what has been achieved previously.

3.2 Sparing Factors

For the purpose of adaptive fractionation, treatment plans and the daily geometric variations are described in terms of sparing factors δ

$$\delta_t = \frac{d_t^{\mathrm{N}}}{d_t}$$

where d_t^{N} denotes the dose received by the dose-limiting OAR in fraction t and d_t the dose delivered to the tumor. Clinical practice of dose prescription and constraint specification is followed for the definition of d_t^{N} and d_t: The dose to the OAR d^{N} is defined as the dose exceeded in 1cc of the OAR ($D_{1\mathrm{cc}}$), which is a commonly used dose parameter for bowel, stomach or duodenum in SBRT treatments. The tumor dose d_t is defined as the dose exceeded in 95% of the PTV volume ($D_{95\%}$), which is a commonly used dose parameter for dose prescription and reporting [17][18]. Each patient is thus described via a sequence of six sparing factors corresponding to the planning MR and the five treatment fractions. For each tumor every OAR was tracked, that was deemed potentially dose-limiting. For applying adaptive fractionation only patients with not changing dose-limiting OAR were analysed. Further, assumed is that inter-fractional motion is random, such that δ is normal distributed with a patient specific mean μ and standard deviation σ

$$\delta_t \sim \mathcal{N}(\mu, \sigma^2)$$

3.3 Clinical Implementation

From a practical perspective, this approach to adaptive fractionation would be implemented by up-scaling or down-scaling the reoptimised treatment plan for that fraction. That is, we assume that the adaptive radiotherapy process consisting of MR imaging, recontouring, deformable image registration and plan reoptimisation is not altered. The only additional step would be a final up-scaling or down-scaling of the fluence without changing the shape of the dose distribution. This corresponds to a renormalisation of the plan, which is in current practice conducted within a narrow range that can be extended to implement adaptive fractionation the treatment plan process.

3.4 Methods

Pérez Haas [1] presented in his paper how the framework of Markov decision process is applied and how the stochastic optimal control problem is formulated. For that states, actions, state transitions and reward functions to describe interaction with the environment are introduced. Depending on the clinical objective that should be accomplished, in the master thesis [19] three optimisation types with different objectives are presented:

1. To treat a tumor where the desired prescription dose cannot be reached as the OAR is too close to the tumor, the goal was set to maximise the cumulative BED delivered to the tumor subject to the constraint on the cumulative OAR BED.

2. In a case where tumor and OAR are farther apart, the prescribed tumor dose can be obtained without risking the overdosage of the OAR. Therefore, the goal is to minimise the cumulative BED delivered to the OAR subject to delivering the prescribed dose to the tumor.

3. Deciding on which algorithm to use at the beginning of a treatment poses a problem, as it is not known what the average distance will be. Hence, an objective has been set, where the goal is to reach the prescribed tumor dose subject to the constraint on the cumulative OAR BED. If the prescribed dose can be reached, the objective is to minimise OAR BED. If the prescribed dose can not be reached, the tumor dose is maximised.

The paper presented results from the optimal policy applied real patient data using some assumptions about the probability distributions. Also described are optimal policy applied to simulated patient data to show the difference of accumulated tumor BED to reference plan.
In this work, the 2. description of OAR BED minimisation will be discussed again with the modifications for adapting the number of fractions. For that states, actions, state transitions and reward functions to describe interaction with the environment are introduced. Additionally, two approaches to update the model of the environment are presented.

3.5 Arithmetical Implementation

In the paper of Pérez Haas [1] a codebase was built to solve the optimal policy and made available to the public. Essentially the code was built around solving the Bellmann Equations introduced in Eq. 5 and Eq. 6. The codebase architecture design follows a functional oriented programming design; the program is constructed by applying and composing functions. Helper functions sample action space, convert between BED and physical dose, fit hyperparameters to patient data and sample probability distributions. Basic functions define states and reward and calculate the optimal policy with a tabular search solving the Bellmann Equation in a backwards recursive manner; the value function in fraction t is dependent on the future value v_{t+1} marginalised over the sparing factor distribution. The basic functions calculate the optimal policy for a single fraction and are called by treatment functions that calculate a complete treatment in retrospective. Treatment functions iterate through basic functions and feed current states, previous sparing factors and current sparing factors to the basic function, while keeping future sparing factors hidden. State, actions, environment model and value functions are described by discrete values. Thus, the policy function is sampled discretely with parameters defining step sizes. In the master thesis of Pérez Haas [19] there were developed multiple basic functions that are different types to calculate the optimal policy.

4 Adapting Number of Fractions

For favourable patient geometries Adaptive Fractionation strategy has shown to deliver small residual doses in final or close-to-final fractions. The rationale for reducing number of fractions in Adaptive Fractionation is to model the usage of an additional fraction in terms of BED, such that one can minimise the number of fractions used for a treatment. The model aims to prevent applying a strategy, where small residual doses are delivered in the final or close-to-final fractions for favourable patient geometries. The concept introduced in this work, is to not only adapt the dose delivered to the tumor; the number of fractions used in the treatment is subject to adaptation as well. The underlying model to adapt number of fractions is based on the approach discussed in Pérez Haas et al. [1]. Conceptually adapting number of fractions is achieved by employing the same paradigm from adaptive fractionation, but to further increase dose on favourable treatment days, to possibly shorten total treatment time. That is increasing the dose on favourable geometries, such that it might be possible to finish the treatment in an earlier fraction than the prescribed maximum number of fractions; and reduce the dose for unfavourable geometries, such that the treatment conforms to the dose modification

presented in [1] where no fractions are omitted for the treatment. This concept demands constiting a model, that steers the trade-off between using fewer number of fractions and optimising dose delivered to OAR compared to uniform treatments.

4.1 Motivation

Prolonged total treatment time often presents a burden to patient and discomfort. For patients in an image guided radiotherapy treatment with plan adaptation this means lying in treatment position for more than half an hour, with visiting time extending to almost an hour in each fraction, when including preparation for positioning. In many cases patients undergo additional cancer treatment modalities other than radiotherapy, further weakening individuals and radiotherapy causing even greater discomfort. If fractions could be omitted while sustaining the same therapeutic ratio to uniform fractionation, treatment time slots are freed which could be filled by other patients. In summary shortening total treatment time could relieve patients of discomfort caused by the treatment session and free treatment allocation slots at the accelerator providing access to the MR-Linac for additional patients.

4.2 Methods

Extended BED Model

It is assumed that the standard BED model can be extended to varying doses per fraction such that at the end of the treatment the cumulative BED is given by the sum of the BED values delivered in individual fractions. Cumulative dose delivered in tumor is thus

$$B_t^{\mathrm{T}} = \sum_{\tau=1}^{t} (d_\tau + \frac{d_\tau^2}{(\alpha/\beta)_{\mathrm{T}}}) \tag{1}$$

where d_τ denotes the dose delivered to the tumor and δ_τ the sparing factor in fraction τ. Consequently, the cumulative BED delivered to the OAR is

$$B_t^{\mathrm{N}} = \sum_{\tau=1}^{t} (\delta_\tau d_\tau + \frac{\delta_\tau^2 d_\tau^2}{(\alpha/\beta)_{\mathrm{N}}}) \tag{2}$$

In this work, the α/β ratios for the OARs and the tumors are the same for all patients and set to $(\alpha/\beta)_{\mathrm{N}} = 3$ and $(\alpha/\beta)_{\mathrm{T}} = 10$ [20]. Correspondingly, the cumulative biological effective doses in the OAR and the tumor will be denoted as BED_3 or B^{N} and BED_{10} B^{T}. Note that the calculation of cumulative BED_3 in Eq. (2) assumes that the same 1cc of the dose-limiting OAR receives the highest dose, which may not be the case in reality. In this case, Eq. (2) can be considered a worst-case measure for OAR dose, which overestimates the cumulative BED_3 received by any part of the OAR. However, due to the impracticality of deformable dose accumulation in the abdomen, the same approximation is done in current clinical practice.

MDP Model

To determine the optimal doses d_t, we apply the framework of Markov decision processes (MDP) and formulate adaptive fractionation as a stochastic optimal control problem. Here, we first describe the MDP model for a known probability distribution $P(\delta_t)$ and afterwards discuss how to estimate and update $P(\delta_t)$. Optimal control problems are described by states, actions, state transitions and reward functions. In this application, these are given by:

State: In each fraction, the state of a patient's treatment is described by a two-dimensional vector $s = (\delta, B)$ that specifies current sparing factor δ and the cumulative BED B that has been delivered so far in previous fractions. Thus, the state of a treatment in fraction t for a patient with sparing factors $\{\delta_\tau\}_{\tau=1}^{t}$ treated with doses $\{d_\tau\}_{\tau=1}^{t-1}$ is

$$s_t = \left(\delta_t, \sum_{\tau=1}^{t-1} (d_\tau + \frac{d_\tau^2}{(\alpha/\beta)_{\mathrm{T}}})\right)$$

Action and policy: The actions correspond to the physical doses d_t that are delivered to the tumor in a fraction. Thus, a policy specifies for each fraction t and possible state of the treatment, the dose that should be delivered in this state. Tumor doses are constrained by a maximum dose per fraction d^{max} and a minimum dose per fraction d^{min}.

State transition: If, in fraction t, the treatment is in state $s_t = (\delta_t, B_{t-1}^T)$ and a dose d_t is delivered to the tumor, the state transitions to

$$s_{t+1} = \left(\delta_{t+1}, B_{t-1}^{\mathrm{T}} + d_t + \frac{d_t^2}{(\alpha/\beta)_{\mathrm{T}}}\right)$$

in fraction $t+1$. The BED-component of the future state is calculated by adding the tumor BED delivered in fraction t to the previously delivered BED B_{t-1}^{T}, which is assumed deterministic (i.e. we don't consider uncertainty in dose delivery). The sparing factor in fraction $t+1$ is random, making the state transition probabilistic. The probability distribution for the state transition is simply given by the probability distribution over the sparing factors, $P(\delta)$.

Immediate Reward: In each fraction t, the immediate reward r_t is given by the BED delivered to the OAR in that fraction

$$r_t(d, \delta, B^{\mathrm{T}}) = -\left(\delta_t d + \frac{\delta^2 d^2}{(\alpha/\beta)_{\mathrm{N}}}\right) - c(d, B^{\mathrm{T}})$$

Where $c(d, B^{\mathrm{T}})$ is a general penalty term, which is a function of the action and accumulated tumor BED state, in which the future prescribed tumor dose is not met. Constructed is a penalty with a numeric value C

$$c(d, B^{\mathrm{T}}) = \begin{cases} C & \text{if } \left(B^{\mathrm{T}} + d + \frac{d^2}{(\alpha/\beta)_{\mathrm{T}}}\right) < B_{\mathrm{pres}}^{\mathrm{T}} \\ 0 & \text{else} \end{cases} \tag{3}$$

This reward model introduces iso-rewarding actions: for a fixed state there exist two actions that yield the same reward. Assuming the dose d^* reaches the prescription dose and yields the reward r^*. There exists a lower dose d^{**} with the same reward. The dose d^{**} acts as a threshold dose, for finishing the treatment in the current fraction: doses equal or larger than the threshold d^{**} and smaller than d^* are never applied. The greater the parameter C the lower this threshold dose will be to reach the fixed r^*. Parameter C acts as an inverse scaling factor for this threshold.

$$C \propto \left(\delta(d^* - d^{**}) + \frac{\delta^2 \left(d^{*2} - d^{**2} \right)}{(\alpha/\beta)_{\mathrm{N}}} \right)$$

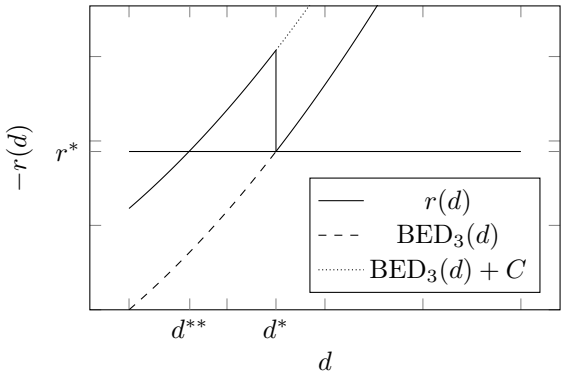

Figure 1: Qualitative description of reward function by composing negative immediate reward from BED_3 without penalty C and BED_3 with penalty term C.

A characteristic of the specified MDP model is that the cumulative BED delivered to the OAR is not part of the state s. It is only integrated in the reward r_t as a penalty. The reason being, that the optimal policy does not depend on the OAR state, i.e. the optimal dose to deliver does not depend on the previously accumulated BED in the OAR. Intuitively it is aimed to minimise future OAR BED, in each state and fraction, independently of the previously accumulated BED in the OAR. However, if the goal was to escalated tumor dose while delivering a fixed OAR BED, the cumulative OAR BED would be part of the state.

Environment Model: The model of the environment provides the probability of arriving in state $s_{t+1} = (\delta_{t+1}, B_t^T)$ with reward r_t, starting from state s_t and choosing action d_t. Arriving at the cumulative tumor BED B_t^T is deterministic, as it is dependent on the previous delivered tumor BED plus action d_t. Arriving at the sparing factor δ_{t+1} however is a stochastic process, where δ is assumed to be normal distributed. Establishing an assumption of the environment model, completely characterises the dynamics of the Markov Process.

Reward Construction

Achieved should be an adaptive fractionation model that accurately steers the expected number of fractions necessary to attain a prescription tumor BED for a treatment. To determine the optimal parameter C a discrete optimisation problem is formulated. To that end a penalty in the form of Eq. (3) with $C = 0$, prescription dose $B^{\mathrm{T}}_{\mathrm{pres}}$ and expected probability distribution $P(\delta)$ is set. Defining a treatment for n number of fractions, calculating the Bellmann Equations with these stipulations and applying the optimal policy to a set of sparing factors $\{\delta_\tau\}_{\tau=1}^n$, yields terminal cumulative BED B_n^{N}. Repeating this process by sampling sets of $\{\delta_\tau\}_{\tau=1}^n$ and averaging B_n^{N} yields an estimate of the average terminal accumulated BED in the OAR written as $\bar{B}_n^{\mathrm{N}}$. This is interpreted as a cost function, seeked to be minimal. In Fig. A.1 the qualitative example with terminal cumulative BED's B_n^{N} can be seen.
To finishing the treatment with prescribed number of fraction n_{pres} the reward function is constructed with a penalty parameter. The penalty parameter is given by C and penalises each additional fraction that is used for the treatment. The cost can thus be extended by linearly adding C for every fraction n to $\bar{B}_n^{\mathrm{N}}$ which yields the total cost function

$$B_n = C \cdot n + \bar{B}_n^{\mathrm{N}}$$

where the numeric value for C can be then evaluated as minimising an objective function dependent on the minimum of B_n

$$C = \underset{c}{\operatorname{argmin}} \left| n_{\mathrm{pres}} - \underset{n}{\operatorname{argmin}} \left[B_n \right] \right|$$

The penalty parameter C can be viewed as a marginal cost with units BED per fraction. Additional fractions are used, so long that reduction of B^{N} is larger than C. Finding the optimal parameter C for a fraction using n_{pres} number of fraction following equation must be fulfilled

$$\bar{B}^{\mathrm{N}}_{n_{\mathrm{pres}}} - \bar{B}^{\mathrm{N}}_{n_{\mathrm{pres}}+1} \geq C$$

Fractionation Decision

To translate the problem into a continuous optimisation problem, the average terminal cumulative BED $\bar{B}_n^{\mathrm{N}}$ will be modeled according to optimal fractionation decision-making. To minimise cumulative OAR BED B_n^{N} for a fixed fraction size d_τ and constant sparing factor with respect to the number of fractions n

$$\min_n \left[B_n^{\mathrm{N}} \right] = \min_n \left[n \delta_\tau d_\tau (1 + \frac{\delta_\tau d_\tau}{(\alpha/\beta)_{\mathrm{N}}}) \right]$$

subject to tumor dose prescription

$$B_n^{\mathrm{T}} = n d_\tau (1 + \frac{d_\tau}{(\alpha/\beta)_{\mathrm{T}}})$$

yields

$$B_n^{\mathrm{N}} = \delta^2 n d_T(n) \left[\frac{1}{\delta} - \frac{(\alpha/\beta)_N}{(\alpha/\beta)_T} \right] + \delta^2 B_n^{\mathrm{T}} \frac{(\alpha/\beta)_N}{(\alpha/\beta)_T}$$

Note that the function is continuous, and the second term is independent on n. Omitting this term and

switching from subscript n to a continous function, the BED $\bar{B}_n^{\mathrm{N}}$ can be fitted to

$$B^{\mathrm{N}}(n;\delta) = \delta^2 n d_T(n) \left[\frac{1}{\delta} - \frac{(\alpha/\beta)_N}{(\alpha/\beta)_T}\right] \tag{4}$$

where d_T is the physical dose

$$d_T(n) = \frac{\sqrt{n(\alpha/\beta)_T(n(\alpha/\beta)_T + 4B^{\mathrm{T}}_{\mathrm{pres}})} - n(\alpha/\beta)_T}{2n}$$

Note that now there is no subscript to B^{N} as it is a continuous function. This allows for a continuous objective function to be minimised for finding C. In the continuous case after fitting $B^{\mathrm{N}}(n;\delta)$, C can be determined analytically

$$C_{n_{\mathrm{pres}}} = \left.\frac{\mathrm{d}B^{\mathrm{N}}(n;\delta)}{\mathrm{d}n}\right|_{n=n_{\mathrm{pres}}}$$

In Fig. 2 for $n_{\mathrm{pres}} = 5$ prescribed number of fractions the parameter $C_5 = 4.36$Gy was determined.

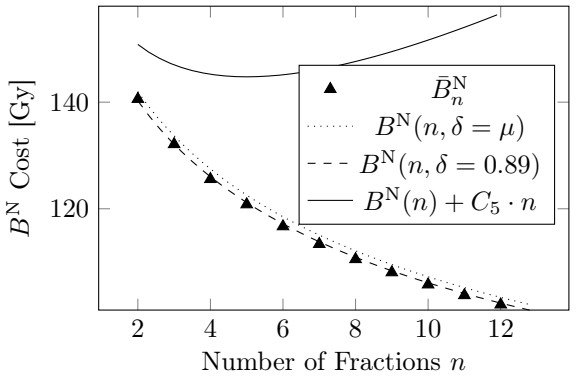

Figure 2: Negative immediate reward or cost B_n^{N} for a gaussian sparing factor distribution $\mu = 0.9$, $\sigma = 0.04$, $B^{\mathrm{T}}_{\mathrm{pres}} = 72$. Sampled are 1000 patients with 5 fractions each that stem from a gaussian distribution with $\mu = 0.9$, $\sigma = 0.04$, $B^{\mathrm{T}}_{\mathrm{pres}} = 72$. Values fitted with $B^{\mathrm{N}}(n;\delta)$.

Also plotted in Fig. 2 is the fitting function $B^{\mathrm{N}}(n;\delta = \mu)$, being the cumulative OAR BED for Uniform fractionation. Cumulative BED is slightly higher compared to $B^{\mathrm{N}}(n;\delta = 0.89)$ fitted to the Adaptive Fraction results. Results stem from simulated treatments, where cumulative BED was averaged for 1000 Patients for each n where δ is the fit parameter. The fit parameter is lower than the distribution sparing mean $\mu = 0.9$.

4.3 Dynamic Programming Algorithm

A dynamic programming (DP) algorithm can be used to compute the optimal policy with the help of a value function [21]. The value function v_t describes how desirable it is to be in state s_t in fraction t and, therefore, it contains the information whether an action should be taken to reach that state. In this application, the value for each state represents the expected cumulative BED that can be delivered to the tumor in the remaining fractions, starting from that state and acting according to the optimal policy. The Bellman equation relates the value function in fraction t to the optimal policy and the value function in the subsequent fraction, which for this application reads

$$v_t(\delta,B^{\mathrm{T}}) = \max_d \Bigg[r_t(d,\delta,B^{\mathrm{T}}) + \ldots \sum_{\delta'} P(\delta') v_{t+1}\left(\delta', B^{\mathrm{T}} + d + \frac{d^2}{(\alpha/\beta)_{\mathrm{T}}}\right)\Bigg] \tag{5}$$

and the policy reads

$$\pi_t(\delta,B^{\mathrm{T}}) = \operatorname*{argmax}_d \Bigg[r_t(d,\delta,B^{\mathrm{T}}) + \ldots \sum_{\delta'} P(\delta') v_{t+1}\left(\delta', B^{\mathrm{T}} + d + \frac{d^2}{(\alpha/\beta)_{\mathrm{T}}}\right)\Bigg] \tag{6}$$

The value function $v_t(\delta, B^{\mathrm{T}})$ is the sum of the immediate reward and the future expected reward, maximised with respect to the dose. The immediate reward consists of reward from immediate OAR BED and weighted reward from not terminating the treatment in the current fraction. The future expected reward is the future value weighted with the sparing factor probability. Abstractly speaking, the value function displays the best possible immediate and future reward summed, the algorithm can secure in each state.

In return the policy describes the optimal strategy to maximise the value function. Value function and optimal policy can be calculated iteratively in one backward recursion starting from the last fraction. To enforce the cumulative tumor BED prescription $B^{\mathrm{T}}_{\mathrm{pres}}$, a terminal reward of -$\infty$ is assigned to all terminal states in which the cumulative tumor BED prescription is not met after the last fraction. To that end, the terminal reward corresponding the value function v_{F+1} at the end of the treatment after all F fractions are delivered, is initialised to

$$v_{F+1}(\delta_{F+1}, B_F^{\mathrm{T}}) = \begin{cases} 0 & \text{if } B_F^{\mathrm{T}} = B^{\mathrm{T}}_{\mathrm{pres}} \\ -\infty & \text{else} \end{cases}$$

This initialises the value function and optimal policy in the last fraction F. Practically speaking as a consequence, optimal policy in the last fraction will simply deliver the maximum residual tumor BED, to end up at the prescribed cumulative tumor BED given in B_5^{T}. Such a policy exploits above value function initialisation equation. In case prescription tumor BED cannot be achieved due to constraints in the action, the terminal reward can be initialised to a linear penalty function dependent on the difference of cumulative BED to prescription

$$v_{F+1}(\delta_{F+1}, B_F^{\mathrm{T}}) = -s \cdot \left|B_F^{\mathrm{T}} - B^{\mathrm{T}}_{\mathrm{pres}}\right|$$

Where $s \gg B^{\mathrm{T}}_{\mathrm{pres}}$ is an arbitrary large number. This ensures that the policy in the last fraction is to apply the dose that minimises the difference between cumulative and prescribed BED. Policy artefacts from not reaching exactly the prescription in the last fraction due to discretisation of the actions can thus be avoided.

4.4 Expected Remaining Number of Fractions

It is difficult to interpret the value function with regard to the weight C, as the value function is composed of OAR BED rewards and reward for terminating treatment. Thus, to make the value function more interpretable, we calculated expected remaining number of fractions in the state space. The remaining number of fractions display the remaining number of fractions, that are expected in every state with. It uses the optimal policy that is already solved and the same assumptions about the probability distribution from the value function. Note that the optimal policy π_t here is in units of BED.

$$\varepsilon_t(\delta,B^{\mathrm{T}}) = \eta_t(\delta, B^{\mathrm{T}}) + \ldots \sum_{\delta'} P(\delta')\varepsilon_{t+1}\left(\delta', B^{\mathrm{T}} + \mathrm{BED}_{10}\left[\pi_t(\delta, B^{\mathrm{T}})\right]\right) \tag{7}$$

where η is a binary function that specifies, if under the optimal policy in the current fraction the prescription dose is achieved or not.

$$\eta_t(\delta, B^{\mathrm{T}}) = \begin{cases} 0 & \text{if } \pi_t(\delta, B^{\mathrm{T}}) \geq B^{\mathrm{T}}_{\mathrm{pres}} - B^{\mathrm{T}} \\ 1 & \text{else} \end{cases} \tag{8}$$

In similar manner to the value function, the terminal remaining number of fractions is initialised to

$$\varepsilon_{F+1}(\delta_{F+1}, B^{\mathrm{T}}_F) = 0$$

as in the last fraction no additional fractions are used.

4.5 Quantification of Benefit

The treatment plans given by Adaptive Fractionation is compared benchmarked with the following treatments

1. A reference treatment in which 5×8 Gy physical dose (14.4 Gy tumor BED, 72 Gy B^{T}_5) is prescribed to the tumor in each fraction. Hence, the reference treatment delivers exactly 72 Gy tumor BED. It is assumed that prescribed tumor BED may be delivered to patients without compromising OAR BED constraint in the dose-limiting OAR.

4.6 Code Repository

The model described in this work was implemented into the existing codebase that can be found in the public repository [22]. Furthermore, the codebase was rebuilt the suit a command line interface, speed up calculation and streamline logic flow. These changes and extensions are given by

Hardcoded Settings: Probability distribution, states and actions are discrete. Parameters defining step sizes, upper- and lower bounds for these variables were previously hardcoded whereas now they can be chosen by the user. So the resolution of the optimal policy can be adapted to a specific problem. The user can also quickly generate plots of optimal policy, value function and expected remaining number of fractions.

Interpolation: It is possible that the optimal policy for reducing number of fraction is discontinuous, as for some states the remaining dose will be delivered. Discretisation of the states and actions led to artefacts (see Fig. A.4) in the discontinuous policy. The solution was to introduce linearly sampled BED action space and converting to physical dose. This allows to waive the use of interpolation. This also presumes that the stepsize for state and action have to be the same. In case they are not the same, e.g. the user wants to save computation time and set a state space lower than action space, the interpolation is used automatically.

Redundancy: Multiple basic functions exist for different types to calculate the optimal policy. Instead of basic functions utilising sharing same helper functions and treatment functions the whole workflow was duplicated for each type. Every type used its own set of helper-, basic- and treatment functions which were essentially duplicates of each other. Duplicates of helper functions were made redundant and a package was streamlined that shared helper functions amongst basic functions.

Architecture: The newer streamlined package offers a single treatment function that loops through the basic functions specified by the user. The user is presented with options for type of optimisation, type of probability distribution updating and options for policy calculation. All these input parameters can be specified in a dictionary inside a `python` script or a `json` file.

Command Line Interface: To utilise the package in the command line, an executable is shipped that is invoked with the parameters in the `json` file. The user can specify whether to log output data during the calculation or plot information, settings of the action, policy etc. and keys for choosing optimisation method, objective and constraints.

Parallelisation: Previously in parts of the basic functions parallel computation for the value function was introduced. However, large components were still missing parallelisation. While streamlining the helper and basic functions, computation of the value function was parallelised, by introducing computation that fully supports vectorised matrix multiplication. Storing the future value function v_{t+1} within each step of the iteration, allows to compute the current value function in parallel over all possible states. Parallelisation was also introduced for sampling and the probability distribution which before used first order loop. The run-time of the algorithm was improved by $8 - 16$ times compared to the previous approach.

It runs in the order of 10^{-2} seconds (order of 10^{0} seconds before).

4.7 Results Real Data

For demonstration purposes a sparing factor sequence from real patients is used to apply the model. A patient who's sparing factors are exemplary to demonstrate the model, was chosen. The requirements are specified as: The sparing factors are large enough (between 0.8 to 1.0) that it represents a relevant case in which clinicians face the risk of compromising tumor BED, but sparing factors show large deviation to lower sparing factors, such that Adpative Fractionation is a promising strategy to mitigate this risk. For Adaptive Fractionation to be an advantageous strategy in this test setup, the sparing factors do not necessarily need large standard deviation. It is sufficient that only few sparing factors are exceptionally low compared to the mean, for Adaptive Fractionation to improve treatment quality. However, in clinical practice it is not possible to estimate whether one outling favourable sparing factor will appear through the course of the treatment, from only knowing the planning sparing factor. Rather one needs to estimate the variation of the sparing factor distribution, where outlying sparing factors are more likely for larger variation. For simplicity the treatment model assumes a fixed variation.

Patient Model: The patient model describes the sparing factor sequence. In this case the sparing factor data are real from patients exported from the MR-Linac treatment planning system. In Tab. 1 three candidates are shown that fulfill above requirements.

Treatment Model: The treatment model is the prescription for the patient. It is composed of a defined total number of fractions that can be used $n_{\max} = 5$, a constant C steering the desired number of fractions $n_{\text{pres}} = 4$ that should be used, tumor dose prescription and an assumption of the sparing factor probability distribution. The variation is chosen in terms of standard deviation $\sigma = 0.10$ and the mean to be the planning sparing factor $\mu = \delta_0$. Planned is a treatment with five fractions with a tumor BED prescription of 72 Gy which corresponds to a physical fraction dose of 8 Gy delivered in 5 fractions. For the parameter C two values are chosen to show how the optimal policy behaves. First $C = 0$ with no intention of reducing number of fractions, and additionally the optimal C is evaluated according to the section about optimal reward. The evaluated C_4 are $2.8, 2.0, 2.2$ for patients $3, 7, 13$ and is applied.

Table 1: Patient candidates for adaptive fractionation. Given is the sparing factor in each fraction.

Patient	$t = 0$	1	2	3	4	5
3	0.72	0.79	0.61	0.83	0.77	0.78
7	0.64	0.66	0.69	0.86	0.66	0.57
13	0.84	0.86	0.95	0.88	0.92	0.84

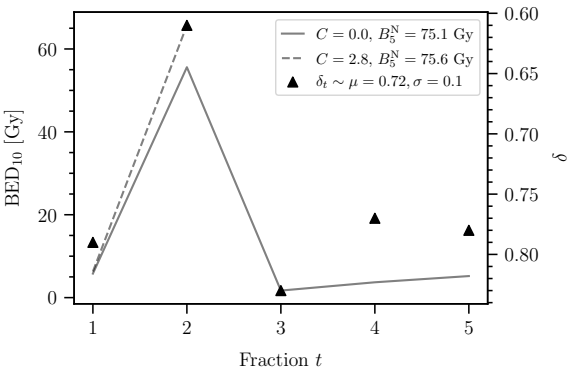

Figure 3: Adaptive fractionated therapy for a 5 fraction treatment applied to patient 3. Normal distribution is estimated from planning sparing factor $\mu = \delta_0 = 0.72$ and fixed $\sigma = 0.1$.

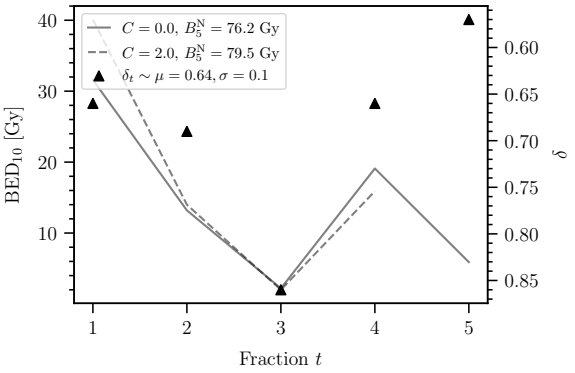

Figure 4: Adaptive fractionated therapy for a 5 fraction treatment applied to patient 7. Normal distribution is estimated from the planning sparing factor $\mu = \delta_0 = 0.64$, $\sigma = 0.1$.

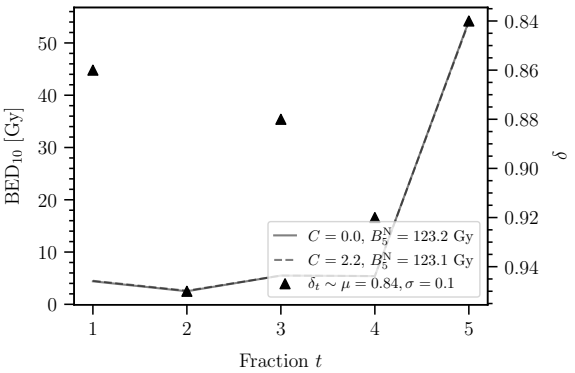

Figure 5: Adaptive fractionated therapy for a 5 fraction treatment applied to patient 13. Normal distribution is estimated from the planning sparing factor $\mu = \delta_0 = 0.84$, $\sigma = 0.1$.

Patient 3: Applying the treatment model to the real sparing factor data of patient 3, the cumulative BED B_5^{N} for $C = 0$ results in 75.1 Gy. Comparing this to Uniform Fractionation, delivering 8 Gy physical dose to the tumor in each fraction yields 91.8 Gy cumulative BED B_5^{N}. Thus, Adaptive Fractionation for this specific patient and treatment model surpasses OAR sparing by delivering 16.7 Gy less cumulative BED B_5^{N} compared to Uniform Fractionation. Applying Adaptive Fractionation with the optimal parameter $C_4 = 2.8$ results in the treatment terminating with 2

Table 2: Comparison of accumulated OAR BED B_5^{N} for Uniform Fractionation (UF) and Adaptive Fractionation (AF). The number of fraction used to complete the tumor prescription dose of 72 Gy is given by n_{frac}.

Patient	Fractionation	B_5^{N} [Gy]	n_{frac}
3	UF	91.8	5
	AF $C = 0$	75.1	5
	AF $C = 2.8$	75.6	2
7	UF	79.0	5
	AF $C = 0$	76.2	5
	AF $C = 2.0$	79.5	4
13	UF	120.3	5
	AF $C = 0$	123.1	5
	AF $C = 2.2$	123.1	5

fractions, while delivering less cumulative BED B_5^{N} than Uniform Fractionation and reaching prescribed tumor dose of 72 Gy. On average, it is expected that the treatment is finished in 4 fractions, but for patient 3 the treatment is terminated after two fractions while accumulating 0.5 BED more compared to Adpative Fractionation $C = 0$. In Tab. 2 the comparison is summarised.

Patient 7: Applying the treatment model to the real sparing factor data of patient 7, the cumulative BED B_5^{N} for $C = 0$ results in 76.2 Gy. Comparing this to Uniform Fractionation, delivering 8 Gy physical dose to the tumor in each fraction yields 79.0 Gy cumulative BED B_5^{N}. Thus, Adaptive Fractionation for this specific patient and treatment model surpasses OAR sparing by delivering 2.8 Gy less cumulative BED B_5^{N} compared to Uniform Fractionation. Applying Adaptive Fractionation with the optimal parameter $C_4 = 2.0$ results in the treatment terminating with 4 fractions, while delivering 0.5 Gy more cumulative BED B_5^{N} than Uniform Fractionation and reaching prescribed tumor dose of 72 Gy. For patient after 4 fractions the treatment is terminated while accumulating 3.3 Gy more BED compared to Adpative Fractionation $C = 0$.

Patient 13: For patient 13 the cumulative BED B_5^{N} for Adaptive Fractionation $C = 0$ results in 120.3 Gy. Comparing this to Uniform Fractionation, delivering 8 Gy physical dose to the tumor in each fraction yields 123.1 Gy cumulative BED B_5^{N}. Thus, Adaptive Fractionation delivers 2.8 Gy more cumulative BED B_5^{N} compared to Uniform Fractionation. Applying Adaptive Fractionation with the optimal parameter $C_4 = 2.0$ gives the same results, while using all 5 fractions.

4.8 Results Synthetic Data

We introduce a synthetic patient is introduced to further demonstrate the model on a sequence of sparing factors. The synthetic patient is a patient not based on real data, but manually specified data. The sparing factors of the synthetic patient are chosen such, that the patient represents a relevant case in which the prescribed tumor BED may be compromised. In this section we will look at the accumulated BED results, when applying the strategies of adaptive fractionation to the synthetic patient. Hence, an according treatment model is composed, that defines prescription to the patient. Demonstrations based on the synthetic patient and treatment model include: calculation of the optimal policy and retrospective application of this policy to the synthetic patient.

Patient Model: In this case the patient model describes a manually chosen sequence of sparing factors for the synthetic patient. Through the retrospective treatment observed sparing factors of this synthetic patient are $\{\delta_\tau\}_{\tau=1}^{t=5} = \{\mu, \mu \pm \sigma, \mu, \mu, \mu\}$. In the second fraction a deviation from the mean sparing factor will appear to demonstrate the applied policy. Note the first sparing factor in fraction $\tau = 0$ is the planning sparing factor and not sampled here, as it is not used. This, the constructed synthetic patient has sparing factor sequence of 0.75 for all fractions except in the second fraction $\delta_2 = 0.65$.

Treatment Model: The treatment model is planned with $n_{\max} = 5$ and 4 desired number of fractions $n_{\mathrm{pres}} = 4$. Planned is the same treatment as with the real patient data, with a tumor BED prescription of 72 Gy which corresponds to a physical fraction dose of 8 Gy delivered in 5 fractions. For the parameter C two values $(0, 1.2)$ are used to demonstrate optimal policy. In addition, the optimal C is evaluated according to the section about optimal reward. The evaluated C_4 to utilise on average 4 fraction is 3.0. The sparing factor probability distribution is assumed to be normal distributed with mean $\mu = 0.75$ and standard deviation $\sigma = 0.10$. A sequence of sparing factors close to 0.75 in every fraction, corresponds to the clinical case, where the tumor dose may be compromised due to the risk of violating OAR dose constraint.

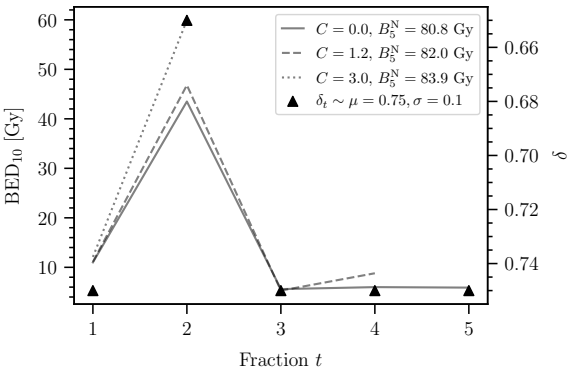

Figure 6: Adaptive fractionated therapy for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$. Synthetic patient has a sparing factor in the second fraction, which is a standard deviation lower than the mean.

Applying the treatment model to the synthetic patient, the cumulative BED B_5^{N} for $C = 0$ results in 80.8 Gy. Comparing this to uniform fractionation (delivering 8

Gy physical dose to the tumor in each fraction) yields an immediate OAR BED of 18 Gy for each fraction with $\delta = 0.75$ and an immediate OAR BED of 14.21 Gy for $\delta_2 = 0.65$, resulting in 86.2 Gy (rounded) cumulative BED B_5^{N}. Thus, Adaptive Fractionation for this specific patient and treatment model surpasses OAR sparing by delivering 5.4 Gy less cumulative BED B_5^{N} compared to uniform fractionation. Applying Adaptive Fractionation with $C = 1.2$ and $C = 3.0$ results in the treatment terminating with 4 respectively 2 fractions (see Fig. 6) while still delivering less cumulative BED B_5^{N} than uniform fractionation and reaching prescribed tumor dose of 72 Gy. In Tab. 3 the comparison is summarised.

Table 3: Comparison of accumulated OAR BED B_5^{N} for Uniform Fractionation (UF) and Adaptive Fractionation (AF). The number of fraction used to complete the tumor prescription dose of 72 Gy is given by n_{frac}. Sparing factors are $\delta_t = 0.75$ except for $\delta_2 = 0.65$

Patient	Fractionation	B_5^{N} [Gy]	n_{frac}
Synthetic	UF	86.2	5
	AF $C = 0$	80.8	5
	AF $C = 1.2$	82.0	4
	AF $C = 3.0$	83.9	2

Value Function

Presented are the corresponding value functions in Fig. 7, 8, 9 and 10. To interpret the results of the value function, the policy and the expected remaining number of fractions for the treatment model are shown.
Optimal policy for the treatment model with $\sigma = 0.1$ and $C = 0$ applied for the synthetic patient can be seen in Fig. 11 and Fig. A.3.

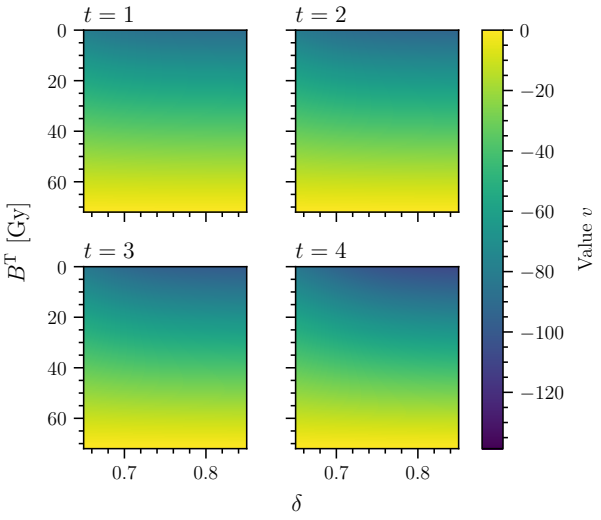

Figure 7: Value function for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$ and $C = 0$.

For parameter $C = 3.0$ the optimal policy is displayed in Fig. 12 and Fig. A.3. In each fraction there is a single plateau visible in the policy function. These plateaux are an area in the state space, that deliver

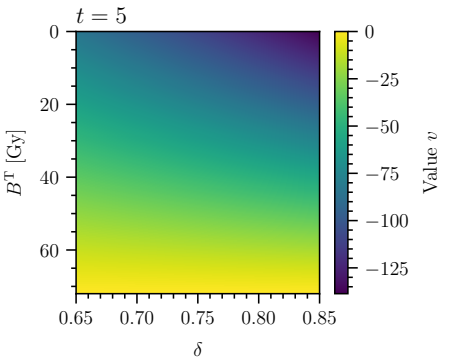

Figure 8: Value function in the last fraction for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$ and arbitrary $C = 0$.

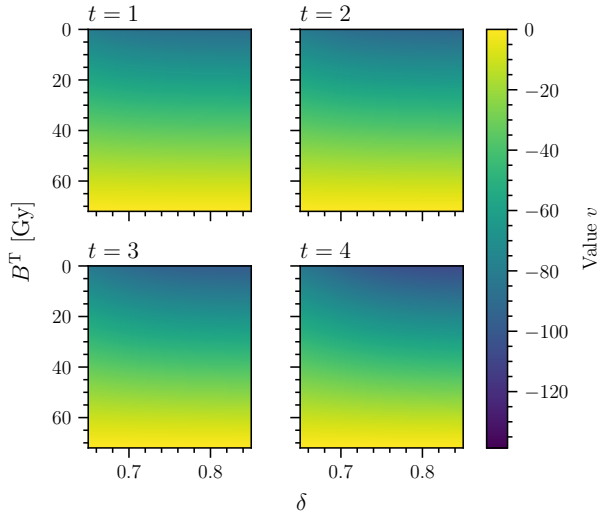

Figure 9: Value function for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$ and $C = 3.0$.

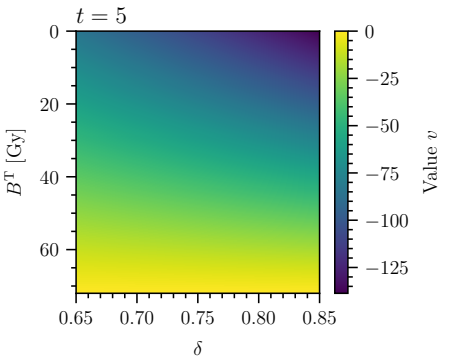

Figure 10: Value function in the last fraction for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$ and arbitrary $C = 0$ or $C = 3.0$.

exactly the remaining dose. Landing on such a plateau in the state space, means the optimal policy is to deliver the remaining dose to finish the treatment in the current fraction.
For the case $C = 0$ and $\sigma = 0.1$ in the expected remaining number of fractions Fig. 14 appears one very narrow plateau on the bottom, for states close to the prescribed tumor dose. The plateau is not visible in the policy. This plateau corresponds to the states where the optimal policy will finish the treatment in

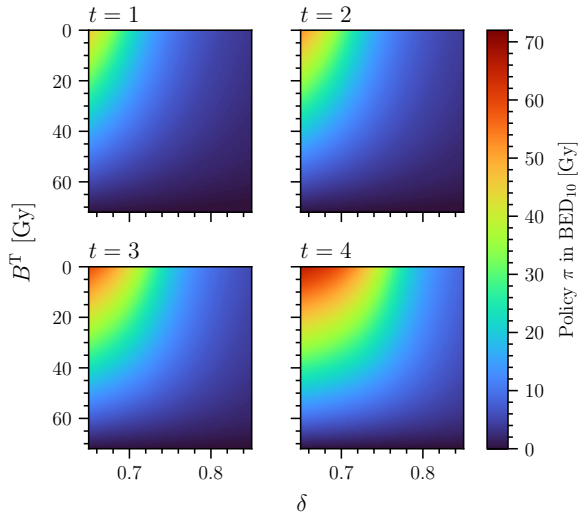

Figure 11: Optimal policy for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$ and $C = 0$. Policy of the last fraction is not shown as it applies the remaining dose regardless of the sparing factor.

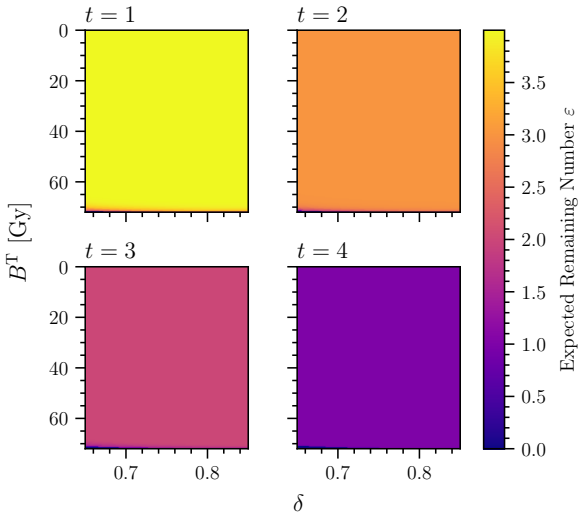

Figure 13: Expected remaining number of fractions for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$ and $C = 0$. Last fraction is not shown as it is zero for every state.

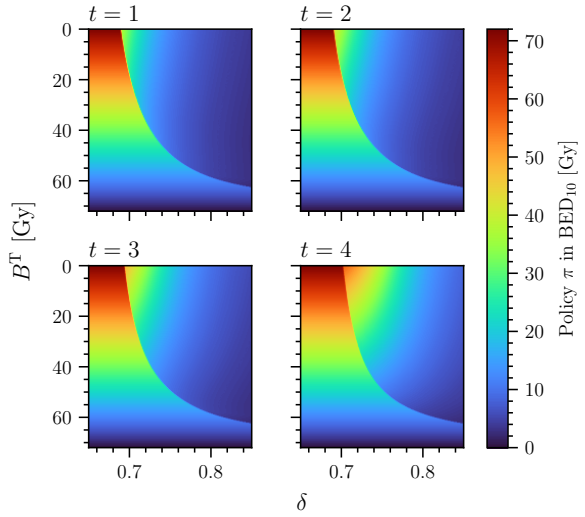

Figure 12: Optimal policy for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$ and $C = 3.0$. Policy of the last fraction is not shown as it applies the remaining dose regardless of the sparing factor.

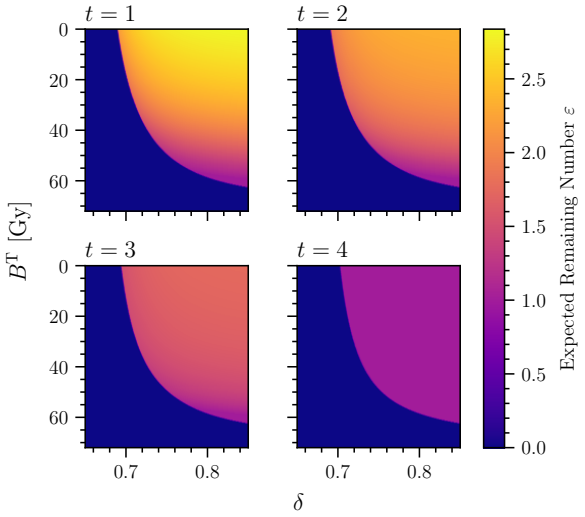

Figure 14: Expected remaining number of fractions for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$ and $C = 3.0$. Last fraction is not shown as it is zero for every state.

the current fraction. Realistically this plateau will not be reached if the optimal dose in the previous fraction is delivered.

In Fig. 14 the case $C = 3.0$, $\sigma = 0.1$ is displayed. Compared to Fig. 13 the plateau in each fraction spans through nearly all the accumulated dose states B^{T}. Comparing the plateaux of the expected remaining number of fractions to the policy they cover the same space. In fraction $t = 4$ the expected remaining number is binary, as the treatment either finishes in the current fraction $t = 4$ or one additional fraction will be used, such that the treatment is finished in fraction $t = 5$. Moving one fraction back to $t = 3$ a plateau is visible again. Between the plateau $\varepsilon = 0$ and $\varepsilon = 2$ there is a gradient where ε takes values between $0, 1$. Such a gradient is mor pronounced in for fraction $t = 1$ and $t = 2$.

5 Patient Data and Treatment Plans

In the results section, real patient data was used to evaluate the model. We looked for patients that are relevant in clinical practice e.g. patients with high sparing factors where the risk of compromising tumor dose is high. And additionally, where sparing factor varies greatly, such that applying Adaptive Fractionation is beneficial for the patient. The show that there exist a benefit for the patient in terms of reducing OAR BED we collected data from patients from three target sites, to look for prospective candidates evaluating the model and discuss if either target sites provide patients where Adaptive Fractionation is beneficial. Considered are patients with abdominal tumors in proximity to either bowel, stomach or duodenum and prostate cancer in proximity to the rectum. For prostate cases urethra and bladder are ignored. These

patients received 5-fraction SBRT treatments at the MR-Linac system (MRIdian, Viewray). All patients were planned and treated according to institutional practice. In addition to the simulation MR and CT scans, daily MR scans were performed for on-line adaptive radiotherapy. Tumors and OARs in a 2cm ring around the tumor were recontoured according to institutional guidelines and daily adaptive treatment plans were created. In each fraction the dose distributions were reoptimised, to adapt to inter-fractional changes, without altering the prescription dose.

5.1 Methods

Patient data was collected manually from the treatment planning system to get an overview of the distance between OAR and tumor volumes in relation to sparing factor. In total 30 patients were extracted from the treatment planning system, 10 for each target. In case of the prostate target 5 patients with CTV as volume of interest and 5 patients with boost target volume as volume of interest were collected. Boost target volumes were collected for patients treated with simultaneous integrated boost strategy [23] [24]. For each patient also extracted were the following variables:

Target: 10 patients were extracted for each tumor that was either pancreas, adrenal glands or prostate cancer. Depending on the predictor the volume for the target was defined for either GTV, PTV, CTV or DIL the volume for prostate boost volume.

OAR: For each patient at least one of the OAR in proximity was extracted. There were two criteria which qualified an OAR to be registered. The first was if the OAR was dose-limiting and the second if the OAR was in close vicinity: Every OAR in one patient that was as close to the tumor as the dose-limiting OAR in at least one fraction, was tracked as well. Only patients were collected, where the dose-limiting OAR did not change during the treatment.

Distance: Distance between target and OAR was the shortest measured distance in all the sagittal, transversal or coronal plane. In principle no overlap was allowed between GTV and OAR. However, if there was an overlap the furthest radial distance of the overlap was measured.

OAR/Target Dose: For pancreas and adrenal glands dose to OAR d^{N} and dose to target d were measured according to the definition in Sec. 3.2. For prostate the parameter definition was different and can be found in Tab. 4 alongside pancreas and adrenal glands parameter definition. The predictors were directly extracted from the DVH in physical dose calculated in the treatment planning system.
A table with which operational quantities for each target were chosen as predictors can be seen in Tab. 4. The predictors were extracted for each tumor-OAR pair for 6 treatment plans in each patient, corresponding to the 5 delivered plans and the initial plan based on the planning MR.

Table 4: Operational quantities collected at the MR-Linac. Explained are target dose definition and volume. Targets are PAN (pancreas), ADN (adrenal glands), prostate (PRO) and BST (prostate boost volume)

Target	w	d^{N} [Gy]	d [Gy]
PAN	GTV	$D_{1\mathrm{cc}}$	D_{95} PTV
ADN	GTV	$D_{1\mathrm{cc}}$	D_{95} PTV
PRO	CTV	$D_{1\mathrm{cc}}$	D_{95} PTV
BST	DIL	$D_{0.1\mathrm{cc}}$	D_{95} DIL

5.2 Patient Data

Pancreas Patients

Sparing factors dependent on spatial distance between target-OAR pair are shown in Fig. 15 for pancreas patients. Sparing factors that were near 1.1 also showed overlap in GTV and OAR. An example of such overlap would be patient 1 for whom sparing factors both with duodenum and stomach are distributed around 1.0 and distance between 0.0 to −0.3 cm. Note that for patient 1 the stomach was the dose-limiting OAR. Information of sparing factors grouped by patients are shown in Fig. 16. Other notable cases are patient 6 and 8 which show only dose-limiting OAR with almost no variation in distance. Patient sparing factors are always closely clustered together. This close clustering for each patient is also visible in Fig. 17, where sparing factors are grouped only by patients without spatial distance. Sparing factors in dose-limiting OAR range from 0.56 to 1.1. Displayed in Fig. 18, is the temporal course of the sparing factors and spatial distances. Tendency is that sparing factor is lower if distance is larger, however not always the case. When looking at patients 6 and 8 where distances are similar, the sparing factor ranges from 1.0 to 1.1 for patient 6 and 0.85 to 1.0 for patient 8.

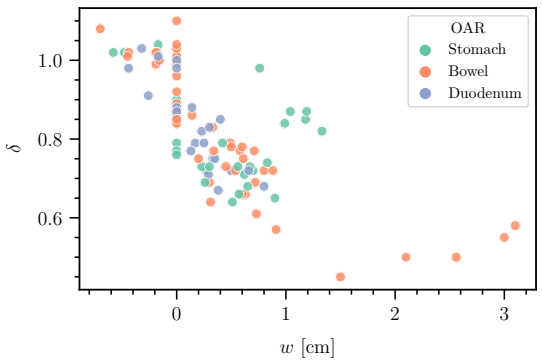

Figure 15: 10 pancreas patients treated at the MR-Linac. Sparing factors δ dependent on distance w grouped by OAR. Multiple OAR were collected per patient and not only the dose-limiting OAR.

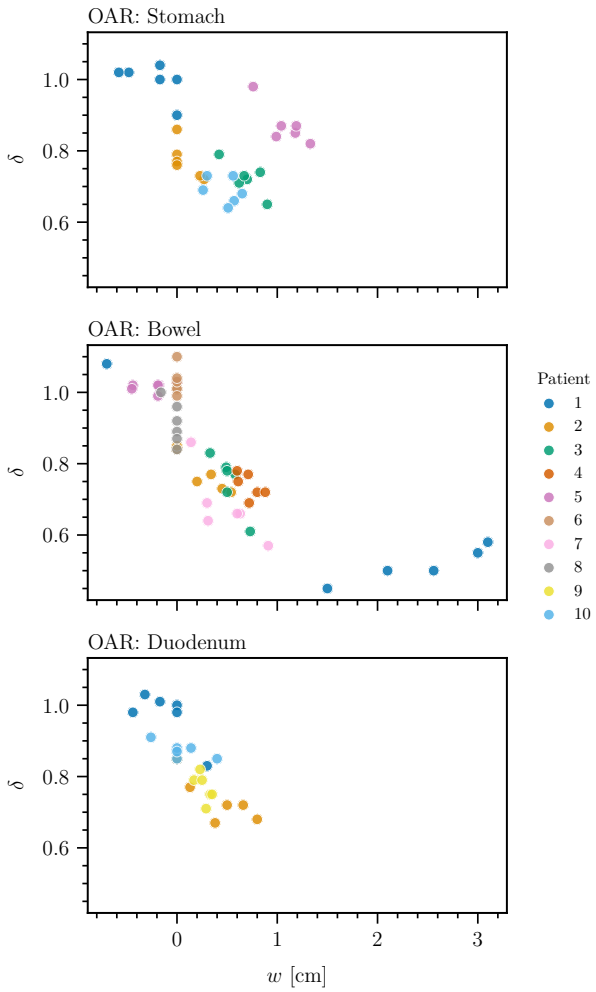

Figure 16: 10 pancreas patients treated at the MR-Linac. Sparing factors δ dependent on distance w grouped by patient. Row shows all collected OAR per patient.

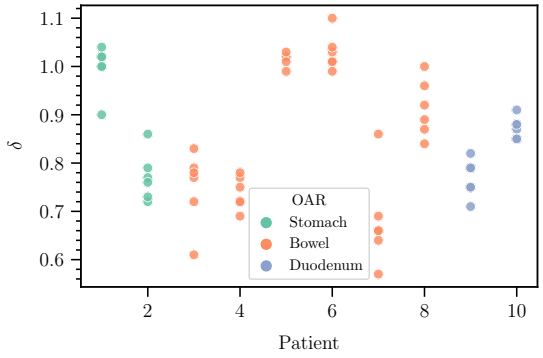

Figure 17: 10 pancreas patients treated at the MR-Linac. Sparing factors δ for each patient grouped by dose-limiting OAR.

Adrenal Glands Patients

Sparing factors dependent on spatial distance between target-OAR pair are shown in Fig. 19 for adrenal glands patients. There are no sparing factors that are larger than 1 and also no overlaps registered in GTV and OAR. Fig. 20 displays the same picture as Fig. 16 but this time additionally grouped by patients. Important to note is that sparing factors range from 0.2 to 1.0 in stomach and 0.05 to 1.0 in bowel. Sparing factors per patient are closely clustered together but

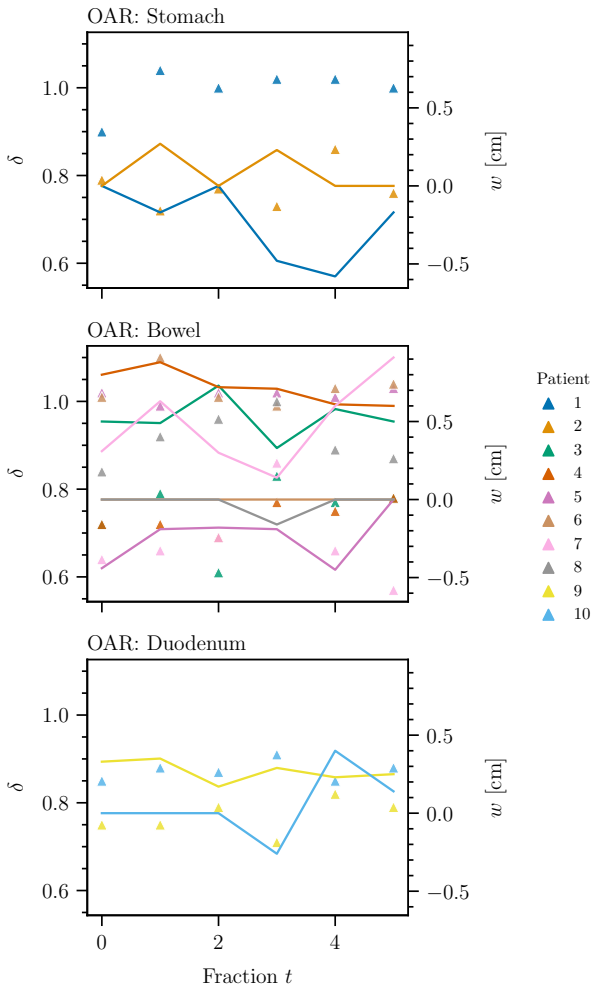

Figure 18: 10 pancreas patients treated at the MR-Linac. Sparing factors δ (shown as triangles) and distance w (shown as lines) displayed for each fraction grouped by patient. Row shows dose-limiting OAR.

across the cohort shows a much wider spread of the variation in sparing factors. Looking at the sparing factors for each patient and only dose-limiting OAR in Fig. 21, it is also visible that they are closer clustered together for each patient compared to Fig. 17. Sparing factors in dose-limiting OAR range from 0.2 to 1.0. Displayed in Fig. 18, is the temporal course of the sparing factors and spatial distances. Tendency is that sparing factor is lower if distance is larger, however not always the case. When looking at patients 6 and 8 where distances are similar, the sparing factor ranges from 1.0 to 1.1 respectively 0.85 to 1.0.

Prostate Patients

Prostate cases are not grouped by OAR but rather by target volume. Only rectum was registered as OAR. There is no variation in spatial distance when looking at prostate CTV in Fig. 23. The volumes always touched but did not overlap in any fraction. Fig. 24 and Fig. 25 shows sparing factors ranging between 0.97 to 1.06 for prostate CTV and 0.89 to 0.97 for boost volume which is lower. Sparing factors between the two target volumes are overall similar in variation, however Fig. 24 displays larger variation in spatial distance in boost volume compared to prostate CTV.

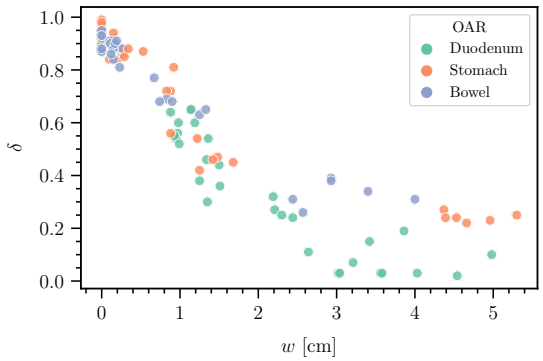

Figure 19: 10 adrenal gland patients treated at the MR-Linac. Sparing factors δ dependent on distance w grouped by OAR. Multiple OAR were collected per patient and not only the dose-limiting OAR.

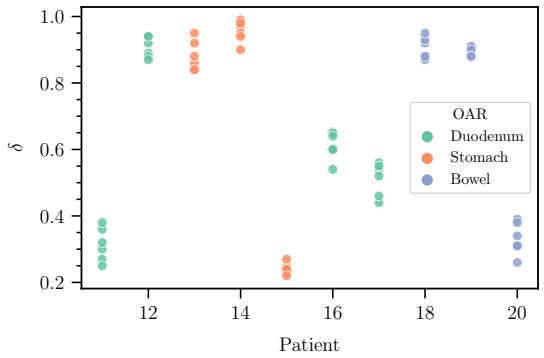

Figure 21: 10 adrenal gland patients treated at the MR-Linac. Sparing factors δ for each patient grouped by dose-limiting OAR.

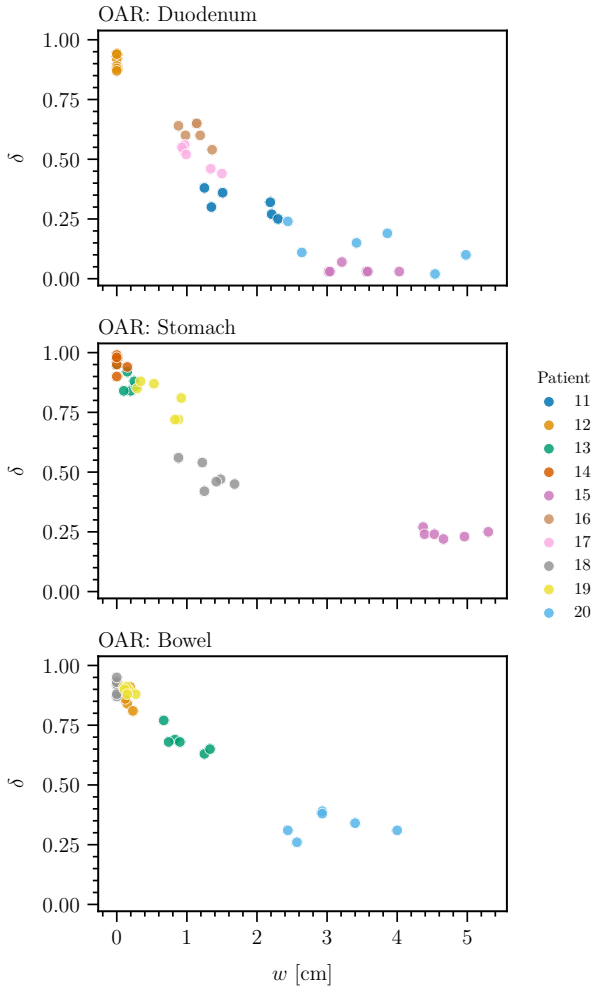

Figure 20: 10 adrenal gland patients treated at the MR-Linac. Sparing factors δ dependent on distance w grouped by patient. Row shows all collected OAR per patient.

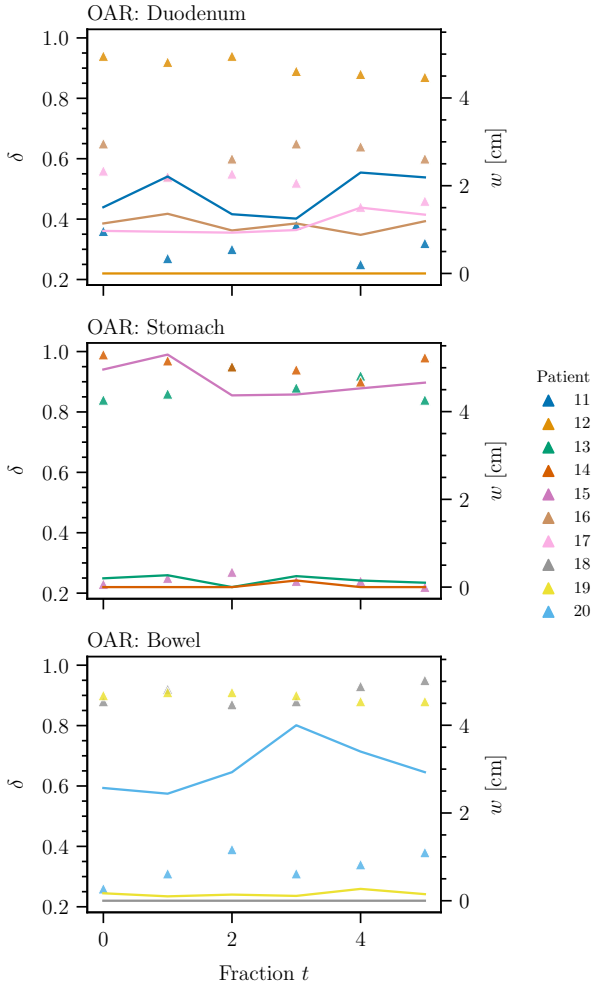

Figure 22: 10 adrenal gland patients treated at the MR-Linac. Sparing factors δ (shown as triangles) and distance w shown (as lines) displayed for each fraction grouped by patient. Row shows dose-limiting OAR.

In Fig. 26 it is visible, that there is no variation in spatial distance for prostate.
The sparing factors of the integrated boost volume are misleading. Dose to the tumor for large distances between rectum and boost volume, was not escalated. Instead, clinicians prescribed and delivered a D_{95} of 40 Gy physical dose in every fraction. For patients 27 and 28 the distance is very large with distance around 1 respectively 1.5cm. In regard to the sparing factor the dose could theoretically have been escalated even further than prescribed D_{95} of 40 Gy physical dose.

6 Discussion

We investigated whether Adaptive Fractionation may improve OAR sparing and reduce number of fractions. Studied were patient cases at risk of compromising tumor dose prescription, due to high sparing factors. If Adaptive Fractionation provided a benefit was measured whether cumulative OAR dose was smaller com-

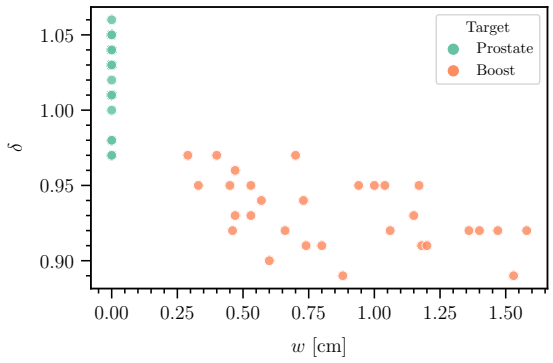

Figure 23: 10 prostate patients treated at the MR-Linac. Sparing factors δ dependent on distance w grouped by target volume BST (prostate boost volume) and PRO (prostate CTV).

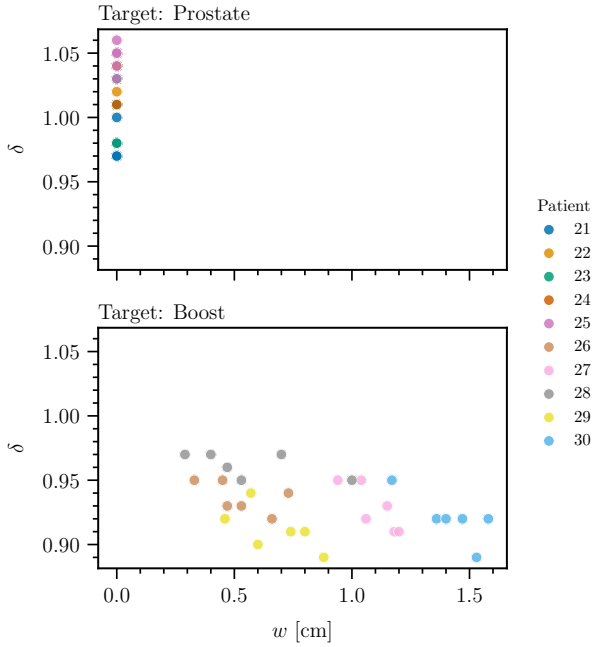

Figure 24: 10 prostate patients treated at the MR-Linac. Sparing factors δ dependent on distance w grouped by target volume BST (prostate boost volume) and PRO (prostate CTV).

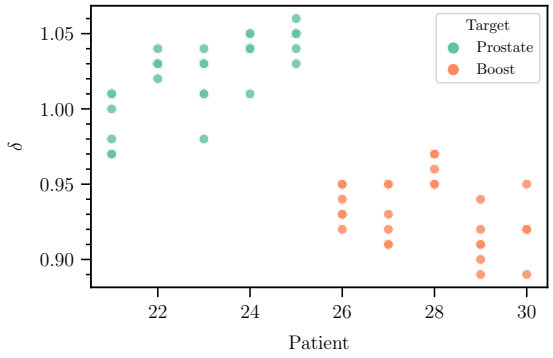

Figure 25: 10 prostate patients treated at the MR-Linac. Sparing factors for each patient grouped by target volume BST (prostate boost volume) and PRO (prostate CTV).

pared to Uniform Fractionation and the treatment used less number of fractions. Two of the three patients showed a benefit in terms of OAR sparing and reducing number of fractions. For patient 13 the OAR BED was higher compared to Uniform Fractionation while also using 5 number of fractions. This can be explained with a planning sparing factor of 0.84 that

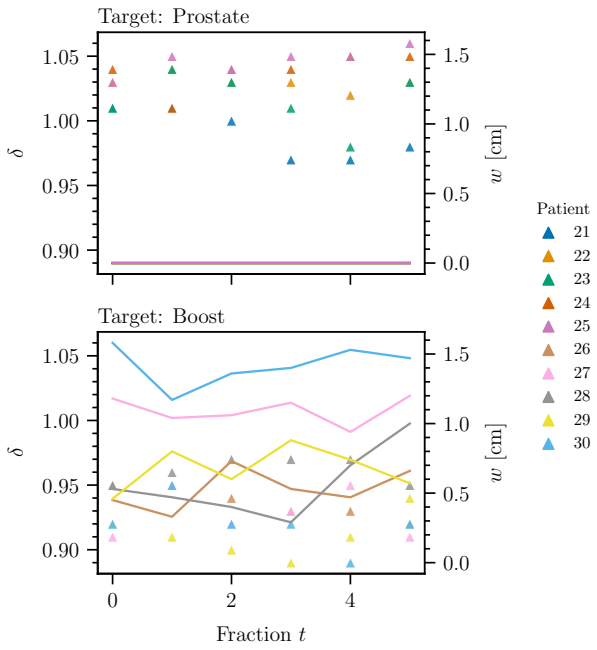

Figure 26: 10 prostate patients treated at the MR-Linac. Sparing factors δ (shown as triangles) and distance w (shown as lines) displayed for each fraction grouped by target volume BST (prostate boost volume) and PRO (prostate CTV)

is not representative of the sparing factor distribution. All sparing factors that appear from fraction 1 to 4 are larger than the expected mean of μ, such that the algorithm hesitates applying large fraction sizes until the last fraction. In the last fraction a sparing factor appears that is marginally smaller than the expected mean.

Prostate patient's sparing factors are not suitable for Adaptive Fractionation, as there is little variation in the sparing factors. Sparing factors from boost volume are not misleading, as clinicians delivered a fixed prescribed dose of D_{90} of 40 Gy regardless of distance to rectum.

7 Conclusion

The Adaptive Fractionation paradigm exploiting day-to-day variations in the distance of the tumor from the dose-limiting OAR was successfully extended to reducing number of fractions. Based on this study considering 5-fraction SBRT treatments of two pancreas tumor patients in proximity to bowel and one adrenal glands tumor in proximity to stomach, it is concluded that the extension brings a benefit to the pancreas patients in terms of OAR dose sparing and reducing number of fractions. However, no other patients show a similar amount of interfraction motion and OAR-to-tumor proximity to substantially benefit from Adaptive Fraction with reduced number of fractions. For the adrenal glands patient Adaptive Fractionation for reducing number of fractions did not bring a benefit to the patient. There is very small amount of inter-fractional motion for prostate patients in proximity to rectum for either CTV or integrated boost volume for Adaptive Fractionation to be beneficial.

References

[1] Y. S. Pérez Haas, R. Ludwig, et al. "Adaptive fractionation at the MR-linac". In: *Physics in Medicine and Biology* (2023). DOI: 10.1088/1361-6560/acafd4.

[2] S. Klüter. "Technical design and concept of a 0.35 T MR-Linac". In: *Clinical and Translational Radiation Oncology* 18 (2019), pp. 98–101. DOI: 10.1016/j.ctro.2019.04.007.

[3] A. Barrett, S. Morris, et al. *Practical Radiotherapy Planning*. 4th ed. CRC Press, 2009. DOI: 10.1201/b13373.

[4] M. Goitein. *Radiation oncology : a physicist's-eye view : Biological and medical physics, biomedical engineering*. Springer New York, 2007. DOI: 10.1007/978-0-387-72645-8.

[5] M. van Herk, P. Remeijer, et al. "The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy". In: *International Journal of Radiation Oncology*Biology*Physics* 47.4 (2000), pp. 1121–1135. DOI: 10.1016/S0360-3016(00)00518-6.

[6] C. Wu, R. Jeraj, et al. "Re-optimization in adaptive radiotherapy". In: *Physics in Medicine and Biology* 47 (2002), pp. 3181–3195. DOI: 10.1088/0031-9155/47/17/309.

[7] D. A. Jaffray. "Image-guided radiotherapy: from current concept to future perspectives". In: *Nature Reviews Clinical Oncology* 9 (2012), pp. 688–699. DOI: 10.1038/nrclinonc.2012.194.

[8] S. Acharya, B. W. Fischer-Valuck, et al. "Online Magnetic Resonance Image Guided Adaptive Radiation Therapy: First Clinical Applications". In: *International Journal of Radiation Oncology*Biology*Physics* 94.2 (2016), pp. 394–403. DOI: 10.1016/j.ijrobp.2015.10.015.

[9] O. Seungjong and K. Siyong. "Deformable image registration in radiation therapy". In: *Radiation Oncology Journal* 35.2 (2017), pp. 101–111. DOI: 10.3857/roj.2017.00325.

[10] D. A. Jaffray, J. H. Siewerdsen, et al. "Flat-panel cone-beam computed tomography for image-guided radiation therapy". In: *International Journal of Radiation Oncology*Biology*Physics* 53.5 (2002), pp. 1337–1349. DOI: 10.1016/S0360-3016(02)02884-5.

[11] J. Lagendijk, B. W. Raaymakers, et al. "MRI/linac integration". In: *Radiotherapy and Oncology* 86.1 (2008), pp. 25–29. DOI: 10.1016/j.radonc.2007.10.034.

[12] M. Guckenberger, J. Wilbert, et al. "Potential of Adaptive Radiotherapy to Escalate the Radiation Dose in Combined Radiochemotherapy for Locally Advanced Non–Small Cell Lung Cancer". In: *International Journal of Radiation Oncology*Biology*Physics* 79.3 (2011). DOI: 10.1016/j.ijrobp.2010.04.050.

[13] B. Jones, R. G. Dale, et al. "The Role of Biologically Effective Dose (BED) in Clinical Oncology". In: *Clinical Oncology* 13.2 (2001), pp. 71–81. DOI: 10.1053/clon.2001.9221.

[14] W. Lu, M. Chen, et al. "Adaptive fractionation therapy: I. Basic concept and strategy". In: *Physics in Medicine and Biology* 53.19 (2008), pp. 5495–5511. DOI: 10.1088/0031-9155/53/19/015.

[15] M. Chen, W. Lu, et al. "Adaptive fractionation therapy: II. Biological effective dose". In: *Physics in Medicine and Biology* 53.19 (2008), pp. 5513–5525. DOI: 10.1088/0031-9155/53/19/016.

[16] J. Ramakrishnan, D. Craft, et al. "A dynamic programming approach to adaptive fractionation". In: *Physics in Medicine and Biology* 57.5 (2012), pp. 1203–1216. DOI: 10.1088/0031-9155/55/5/1203.

[17] ICRU. "Prescribing, recording, and reporting photon beamtherapy". In: *ICRU report Volume 50* (1993).

[18] ICRU. "Prescribing, recording, and reporting photon beamtherapy (supplement to ICRU report 50)". In: *ICRU report Volume 62* (1999).

[19] Y. S. Pérez Haas. "Adaptive fractionation at the MR-linac". MA thesis. University of Zurich, Switzerland, 2022.

[20] C. M. van Leeuwen, A. L. Oei, et al. "The alfa and beta of tumours: a review of parameters of the linear-quadratic model, derived from clinical radiotherapy studies". In: *Radiation Oncology* 13.1 (2018), p. 96. DOI: 10.1186/s13014-018-1040-z.

[21] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: A Bradford Book, 2018.

[22] J. T. Weber, Y. S. Pérez Haas, and R. Ludwig. *Adaptive Fractionation Python Package*. https://github.com/openAFT/adaptfx. 2023.

[23] E. Orlandi, M. Palazzi, et al. "Radiobiological basis and clinical results of the simultaneous integrated boost (SIB) in intensity modulated radiotherapy (IMRT) for head and neck cancer: A review". In: *Critical Reviews in Oncology/Hematology* 73.2 (2010), pp. 111–125. DOI: 10.1016/j.critrevonc.2009.03.003.

[24] Q. Wu, R. Mohan, et al. "'Simultaneous integrated boost' (SIB) IMRT of advanced head and neck squamous cell carcinomas - dosimetric analysis". In: *International Journal of Radiation Oncology*Biology*Physics* 51.3, Supplement 1 (2001), pp. 180–181. DOI: 10.1016/S0360-3016(01)02151-4.

Appendix

A Other Notable Results

Described and illustrated are the other treatment model regarding policy, value function and expected remaining number of fractions.

Treatment Model

For the parameter C multiple values are chosen to show how the optimal policy behaves. The sparing factor probability distribution is assumed to be normal distributed with mean $\mu = 0.75$ and standard deviation $\sigma = \{0.001, 0.05, 0.10, 0.15\}$. A sequence of sparing factors close to 0.75 in every fraction, corresponds to the clinically interesting case, where the tumor dose may be compromised due to the risk of violating OAR dose constraint. Provided enough variation in such a case Adaptive Fractionation could bring a benefit to patients, by reducing cumulative dose to OAR.

Policy

The optimal policy $\pi_t(\delta, B^{\mathrm{T}})$ for a 5 fraction treatment with normal distribution $\mu = 0.75$ and very low standard variation of $\sigma = 0.001$ is given in Fig. A.1 and Fig. A.3. The probability distribution $P(\delta)$ is fixed: i.e. the optimal policy for the entire treatment is solved without updating the probability distribution with progressing treatment.

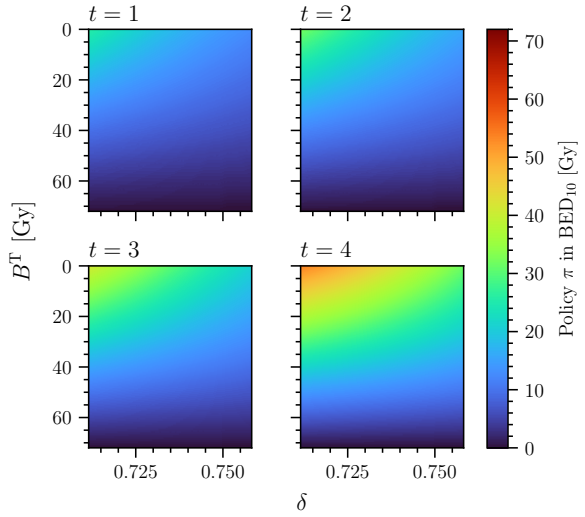

Figure A.1: Optimal policy for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.001$ and $C = 0$. Policy of the last fraction is not shown as it applies the remaining dose regardless of the sparing factor.

For the same expected probability distribution but increasing C gives the policy seen in Fig. A.2 and A.3. A feature of this policy is the appearance of plateaux and thus a discontinuous policy function.

Expected Remaining Number of Fractions

Since the policy is known for every state the expected remaining number of fractions can be calculated for every fraction. In Fig. A.7 the remaining number of fractions are shown to the policy in Fig. A.2 shown before. The bottom plateau in the policy correspond

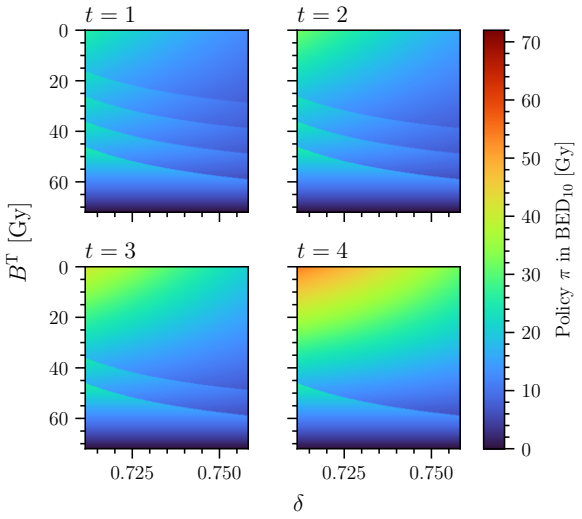

Figure A.2: Optimal policy for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.001$ and $C = 1.9$. Policy of the last fraction is not shown as it applies the remaining dose regardless of the sparing factor.

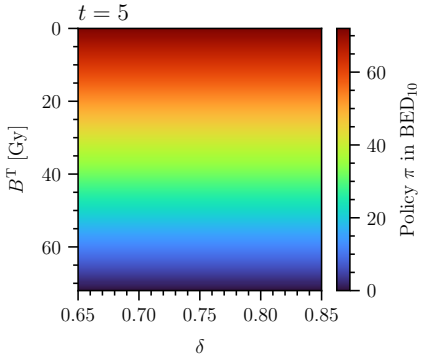

Figure A.3: Optimal policy in the last fraction for an arbitrary 5 fraction treatment. Policy in the last fraction applies the remaining dose regardless of the sparing factor.

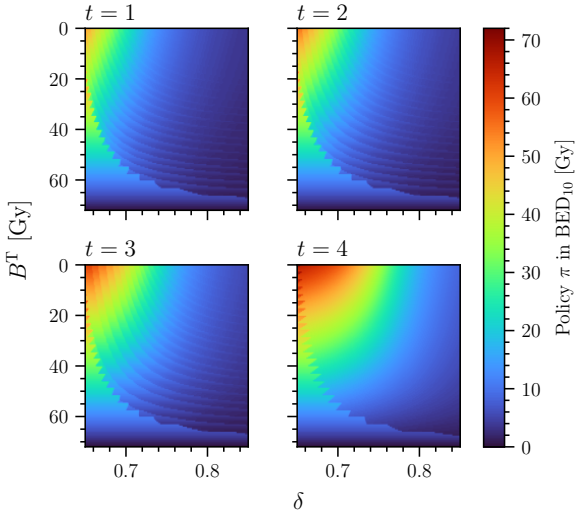

Figure A.4: Optimal policy for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$ and $C = 1.2$, polluted with artefacts due to interpolation.

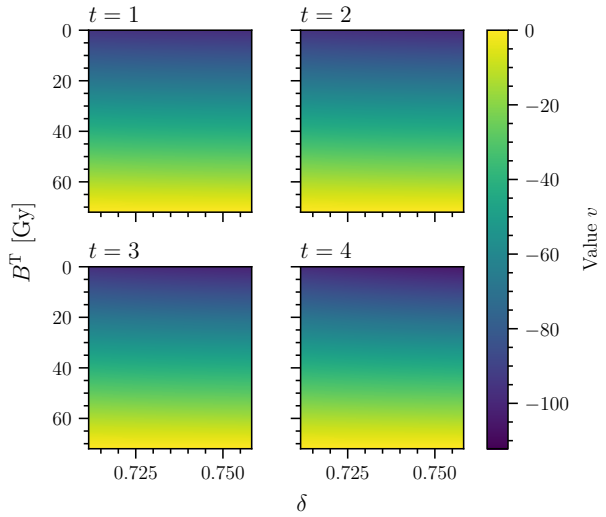

Figure A.5: Value function for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.001$ and $C = 1.9$.

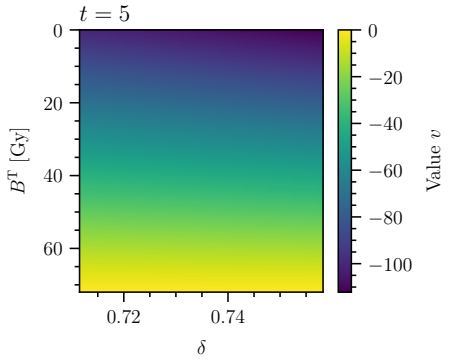

Figure A.6: Value function in the last fraction for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.001$ and arbitrary C.

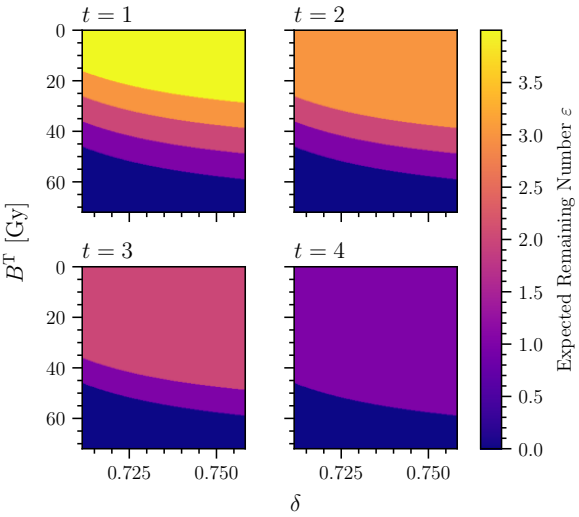

Figure A.7: Expected remaining number of fractions for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.001$ and $C = 1.9$. Last fraction is not shown as it is zero for every state.

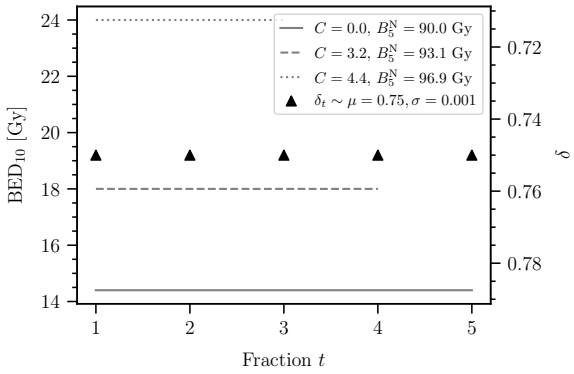

Figure A.8: Adaptive fractionated therapy for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.001$. Sparing factor is the same for every fraction including planning.

to the states where the treatment is will be finished in the current fraction. One plateau above correspond to all states where it is expected to finish treatment the next fraction. The further along in the treatment, the fewer fractions are remaining and thus plateaux disappear. The number of plateaux in a fraction is equal to the number of fractions remaining including the current fraction.
Described and illustrated are the treatment models applied to a synthetic patient with varying observed sparing factors chosen for demonstration.

Constant Sparing Factor

If a synthetic patient shows no deviation at all with sparing factors appearing at $\mu = 0.75$. A uniformly fractionated treatment plan yields the limiting 90 Gy cumulative OAR BED B_5^{N} for a tumor prescription dose of 72 Gy. Assuming an Adaptive Fractionation treatment model with a low variation in sparing factors, such that practically there is no variation expected at all and a constant $C = 0$ (see Fig. A.8), the policy conforms to uniform fractionation. Increasing C yields uniform fractionation policy using fewer fractions and resulting in higher B_5^{N}.
Increasing the expected variation in sparing factors to a standard deviation of $\sigma = 0.05$ seen in Fig. A.9 for otherwise equal treatment model, yields higher cumulative BED B_5^{N} compared to a $\sigma = 0.001$. For $C = 0$ less dose is applied in the first three fractions compared to the treatment model with $\sigma = 0.001$, as a variation in favour of lower sparing factors is expected. As no lower sparing factors appear the dose is steadily increased. Looking at the treatment model with $C = 4.4$ the treatment is finished after 4 fractions with a cumulative BED B_5^{N} 3.5 Gy higher compared to uniform fractionation.
Increasing the expected variation in sparing factors even further to a standard deviation of $\sigma = 0.1$ and $\sigma = 0.15$ seen in Fig. A.10 and Fig. A.11 for otherwise equal treatment model for $C = 4.4$ and $C = 5.0$ parameters only one fraction is omitted.
For a patient model with zero variation in sparing factor and treatment model with large expected variation the optimal policy yields a suboptimal dose delivery reflected by the higher cumulative BED B_5^{N} compared to a treatment model that assumes low variation.

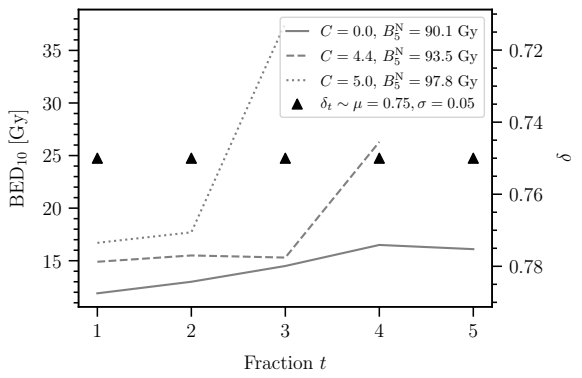

Figure A.9: Adaptive fractionated therapy for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.05$. Sparing factor is the same for every fraction including planning.

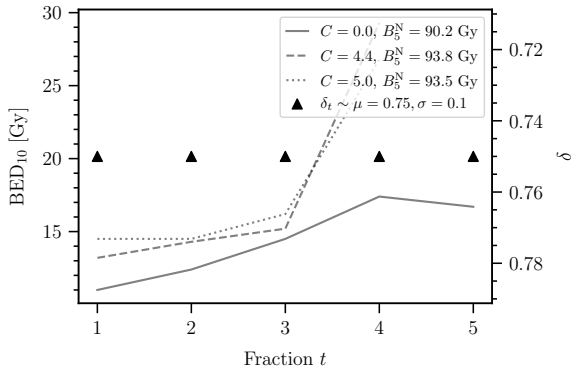

Figure A.10: Adaptive fractionated therapy for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$. Sparing factor is the same for every fraction including planning.

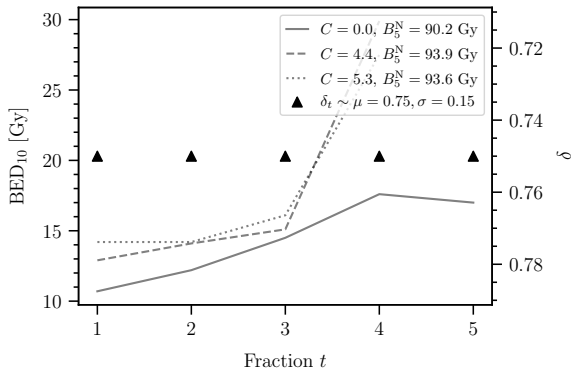

Figure A.11: Adaptive fractionated therapy for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.15$. Sparing factor is the same for every fraction including planning.

Deviation from the Sparing Factor

Applying a treatment model with an expected standard deviation $\sigma = 0.1$ to a patient model with a sparing factor sequence of 0.75 for all fractions except in the second fraction $\delta_2 = 0.85$, the cumulative BED B_5^N compared to a patient model with no variation is lower. The uniform fractionation (delivering 8 Gy physical dose to the tumor in each fraction) yields an immediate OAR BED of 18 Gy for each fraction with sparing factor 0.75 and an immediate OAR BED of 22.21 Gy for $\delta_2 = 0.85$, resulting in 94.21 Gy. Adaptive Fractionation for 5 fraction treatment ($C = 0$) surpasses OAR sparing by delivering 1.81 Gy less cumulative BED B_5^N compared to uniform fractionation.

Adaptive Fractionation with a treatment model of $C = 1.2$ and $C = 3$ do not finish the treatment earlier than 5 fractions (see Fig. A.12) and deliver the same cumulative BED B_5^N as the treatment model with $C = 0$. Nevertheless, the cumulative BED B_5^N is still lower than uniform fractionation while reaching prescribed tumor dose of 72 Gy. The optimal policy does not allow increasing dose to omit fractions, due to the exceptional high sparing factor of 0.85. In Tab. 5 the comparison is summarised.

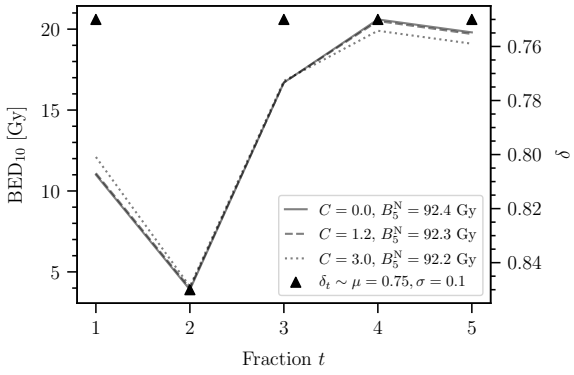

Figure A.12: Adaptive fractionated therapy for a 5 fraction treatment with assumed normal distribution $\mu = 0.75$, $\sigma = 0.1$. Sparing factor in the second fraction is a standard deviation higher from the mean.

Table 5: Comparison of accumulated OAR BED B_5^N for Uniform Fractionation (UF) and Adaptive Fractionation (AF). The number of fraction used to complete the tumor prescription dose of 72 Gy is given by n_{frac}. Sparing factors are $\delta_t = 0.75$ except for $\delta_2 = 0.85$

Fractionation	B_5^N [Gy]	n_{frac}
UF	94.2	5
AF $C = 0$	92.4	5
AF $C = 1.2$	92.3	5
AF $C = 1.3$	92.3	5

B Probability Updating

The DP algorithm relies on a description of the environment to compute an optimal policy, in this case the probability distribution of the sparing factor $P(\delta)$, which we assume to be a Gaussian distribution truncated at 0, with patient-specific parameters for mean and standard deviation. At the start of a treatment, only two sparing factors are available for that patient, from the planning scan and the first fraction. In each fraction, an additional sparing factor is measured, which can be used to calculate updated estimates μ_t and σ_t for mean and standard deviation, respectively.

Maximum a Posteriori Estimation

In each fraction t, the maximum likelihood estimator of the mean of the sparing factor distribution is

$$\mu_t = \frac{1}{t+1}\sum_{\tau=0}^{t}\delta_\tau \tag{9}$$

where δ_0 denotes the sparing factor from the planning MR. The estimator for the standard deviation, given the patient-specific sparing factors up to fraction t, follows a chi-squared distribution, and the maximum likelihood estimator is

$$\sigma_t^{pat} = \sqrt{\frac{1}{t+1}\sum_{\tau=0}^{t}\left(\delta_\tau - \mu_t\right)^2} \tag{10}$$

However, the standard deviation may be severely under- or overestimated if calculated from only two samples at the very beginning of the treatment. Therefore, we assume a population based prior for the standard deviation and compute the maximum a posterior estimator of σ_t via Bayesian inference. As the sparing factors are assumed to follow a normal distribution with unknown variance, a gamma distribution is chosen as prior to estimate the standard deviation σ,

$$f(\sigma; k, \theta) = \frac{1}{\Gamma(k)\theta^k}\sigma^{k-1}\exp\left(\frac{-\sigma}{\theta}\right) \tag{11}$$

with shape k and scale θ the hyperparameters. The maximum a posterior estimator for the standard deviation in fraction t is then

$$\sigma_t = \underset{\sigma}{\operatorname{argmax}}\left[\frac{\sigma^{k-1}}{\sigma^{t-1}}\exp\left(\tfrac{-\sigma}{\theta}\right)\exp\left(\frac{-\left(\sigma_t^{pat}\right)^2}{2\frac{\sigma^2}{t}}\right)\right] \tag{12}$$

Using Eq. (9) and Eq. (12), the probability distribution $P(\delta; \mu_t, \sigma_t)$ is updated with every newly acquired sparing factor and used in the Bellman Eq. (5) and Eq. (6) to recompute the optimal policy before each fraction.

Posterior Predictive Distribution

To predict the distribution of an unobserved sparing factor, a full Bayesian approach is employed. The approach estimates a posterior predictive distribution by marginalising the posterior over the standard deviation.

$$P(\delta; \tilde{\delta}) = \int p(\delta; \mu, \sigma) f(\sigma; \tilde{\delta}) d\sigma$$

The likelihood is defined as

$$p\left(\delta; \mu, \sigma\right) = \left(2\pi\sigma^2\right)^{-t/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{\tau=1}^{t}\left(\delta_\tau - \mu\right)^2\right\}$$

and the constructed conjugate prior is an inverse-gamma distribution:

$$f(\sigma^2; k, \theta) = \frac{\theta^k}{\Gamma(k)}(1/\sigma^2)^{k+1}\exp\left(\frac{-\theta}{\sigma^2}\right) \tag{13}$$

It should be noted that shape k and scale θ too are hyperparameters, as the underlying model parameter is σ. Updating the posterior hyperparameters k_t and θ_t in fraction t with the given sparing factors (including the sparing factor from the planning treatment plan) $\tilde{\delta} = \{\delta_\tau\}_{\tau=0}^{t}$ yields

$$k_t = k + \frac{t}{2} \tag{14}$$

$$\theta_t = \theta + \frac{1}{2}\sum_{\tau=0}^{t}(\delta_\tau - \mu)^2 \tag{15}$$

Resulting as the posterior predictive distribution will be a student t-distribution

$$t_\nu(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{x^2}{\nu}\right)^{-(\nu+1)/2}$$

where the subscript ν refers to the degrees of freedom. The above probability density function is in the standardised form. With parameters μ and σ it can be shifted and scaled. Specifically

$$t_\nu(y; \mu, \sigma) = \frac{t_\nu\left(x = \frac{y-\mu}{\sigma}\right)}{\sigma}$$

Applied to the posterior predictive distribution yields

$$P(\delta; \tilde{\delta}) = t_{2k_t}(\delta; \mu = \mu_t, \sigma = \sqrt{\theta_t/k_t}) \tag{16}$$

Using Eq. (13), (14) and (15), the probability distribution in Eq. (16) is updated with every newly acquired sparing factor and used in the Bellman Eq. (5) and (6) to recompute the optimal policy before each fraction.

C Pseudocode

Algorithm 1 Number of fraction minimisation

Ensure: $T \leq t \leq F$ ▷ *T is the fraction for which optimal policy shall be found*
  for $t = F$ to T **do** ▷ *starting from the last loop through fractions*
    if $B_t^{\mathrm{T}} \geq B_{\mathrm{pres}}^{\mathrm{T}}$ **then**
      break
    else if $t = F$ **then**
      $B_{\mathrm{res}} \leftarrow B_{\mathrm{pres}}^{\mathrm{T}} - B^{\mathrm{T}}$ ▷ *residual BED to reach prescription dose*
      $d_{\mathrm{res}} \leftarrow d \mid \mathrm{BED}_{10}(d) = B_{\mathrm{res}}$ ▷ *respective physical dose to residual BED*
      $v_t\left(\delta, B^{\mathrm{T}}\right) \leftarrow -\mathrm{BED}_3\left(d_{\mathrm{res}}, \delta\right)$ ▷ *penalty from corresponding OAR dose*
      $\pi_t\left(\delta, B^{\mathrm{T}}\right) \leftarrow d_{\mathrm{res}}$ ▷ *optimal policy is residual dose*
    else
      $V_{t+1}\left(d, B^{\mathrm{T}}\right) = \sum_{\delta'} P(\delta') \cdot v_{t+1}\left(\delta', B^{\mathrm{T}} + \mathrm{BED}_{10}(d)\right)$ ▷ *marginalised future value function*
      $r_t\left(d, B^{\mathrm{T}}\right) = -\mathrm{BED}_3(d, \delta) - c(d, B^{\mathrm{T}})$ ▷ *c is penalty for not reaching prescription*
      $v_t\left(\delta, B^{\mathrm{T}}\right) \leftarrow \max_d \left[r_t(d, B^{\mathrm{T}}) + V_{t+1}(d, B^{\mathrm{T}})\right]$ ▷ *Bellmann equation*
      $\pi_t\left(\delta, B^{\mathrm{T}}\right) \leftarrow \mathrm{argmax}_d \left[r_t(d, B^{\mathrm{T}}) + V_{t+1}(d, B^{\mathrm{T}})\right]$

D Patient Data