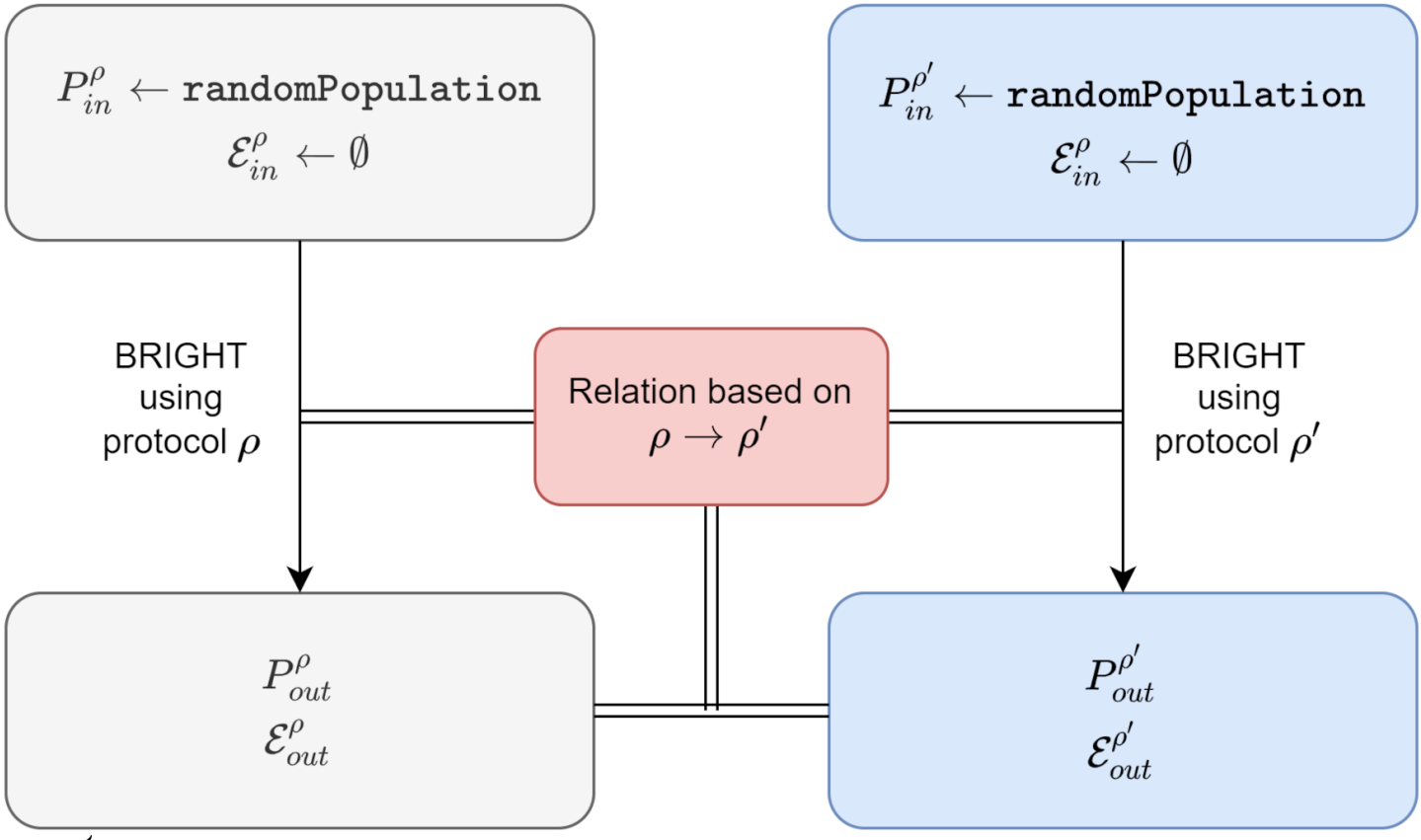
Warm-starting evolutionary plan optimization for high-dose-rate brachytherapy treatment to reduce optimization time

M.C.H. Bakker

MSc. thesis



TU Delft Delft University of Technology

CWI

Warm-starting evolutionary plan optimization for high-dose-rate brachytherapy treatment to reduce optimization time

by

M.C.H. Bakker

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on 23th of June 2022 at 10:00AM.

Student number:	4607654	
Project duration:	February, 2021 – June, 2022	
Thesis committee:	Prof. dr. P.A.N. Bosman,	TU Delft Algorithmics, supervisor
	Prof. dr. M. Specht,	TU Delft Web Information Systems
	Dr. T. Alderliesten,	LUMC Radiation Oncology
	A. Bouter MSc.,	CWI Life Sciences And Health

TU Delft

Warm-starting evolutionary plan optimization for high-dose-rate brachytherapy treatment to reduce optimization time

M.C.H. Bakker

June 16, 2022

Preface

This master thesis concludes the years which I spent at the TU Delft. During my time at the TU Delft, I did learn a lot and it formed me both as a programmer and also as the person I am now. Over the years I got the opportunity to apply my knowledge in different fields. Examples of the different projects I had the pleasure of doing are building software for a pop-up escape room company, building a weeding robot in a multi-disciplinary team, and a cyber-security project in which the goal was to silently communicate with a voice-assistant using ultrasonic sounds. Besides learning a lot, I also had a lot of fun during both my BSc. and MSc. studies.

During my MSc. the course Evolutionary Algorithms (which was taught by Prof. dr. P.A.N. Bosman) sparked my interest, also because of the direct practical applications of evolutionary algorithms which were presented during the course. Because of that, I quickly realized that I did want to do my master thesis research in this direction. I was happy to find out when Prof. dr. P.A.N. Bosman informed me about a possible research topic for a master thesis that is now the subject of my master thesis.

That is where the journey of my master thesis research started. Unfortunately, due to the Covid-19 pandemic, there were no in-real-life meetings. However, that did not prevent me from receiving the guidance I needed during my research from Prof. dr. P.A.N. Bosman, Dr. T. Alderliesten and A. Bouter MSc.

I would like to thank dr. P.A.N. Bosman for the opportunity be able to contribute to the amazing BRIGHT project. I did specifically ask whether dr. P.A.N. Bosman had a project which was more practical than it was theoretical, which this certainly was. I would also like to thank dr. P.A.N. Bosman for the guidance of my research and the tips and feedback he gave during the meetings.

When I saw the feedback of Dr. T. Alderliesten on my draft version of my thesis it was like I saw a colorful palette. After looking at it again I realized that it was all hand-written feedback in all kinds of bright colors which made processing the draft thesis feedback a lot less boring. Besides that, I am also very thankful to Dr. T. Alderliesten for all the feedback which I received during the meetings and also for giving me the opportunity to directly communicate with the radiologists during the KWF BT project team meetings.

To A. Bouter MSc., I would like to thank you for all the feedback and support you provided. No matter what my question was, whether it was a bug in the code or an unexpected result, or whether it was a question about how to tackle certain problems which blocked my progress, you were always available and ready to help. Without all of your help, my research would have been a lot harder.

Finally, I would like to thank everyone, especially my family and friends, who supported me and helped me throughout my journey at the TU Delft. After 5-6 years of studying, I am ready for my next step, which is to make an impact in the Dutch insurance sector using everything I learned during my time at the TU Delft.

Matthias Cornelis Hendrik Bakker
Ottoland, June 2022

Summary

For high-dose-rate brachytherapy, a treatment plan is used to apply sufficient radiation to a tumor while also aiming to avoid irradiating surrounding organs too much. Finding treatment plans is challenging since there is a clear trade-off between target coverage and healthy tissue sparing. To tackle the high-dose-rate brachytherapy planning problem for prostate cancer, BRIGHT can be used to generate a set of plans, which is the application of MO-RV-GOMEA on the high-dose-rate brachytherapy planning problem. BRIGHT has proven to produce high-quality treatment plans which are similar or better in quality than clinical plans which were used in the past which is concluded by an observer study. The set of plans trade-off between target coverage and healthy tissue sparing.

BRIGHT provides plans based on a protocol. This protocol consists of a set of targets that BRIGHT tries to achieve. There are two types of targets within the protocol, namely coverage targets which are aimed at applying a certain amount of radiation to subvolumes that contain cancer cells, and organs at risk (sparing targets) which aim to reduce the radiation applied to surrounding organs. This makes the objective space of brachytherapy planning a multi-objective space. In general, when the coverage of a plan is improved, the sparing will be reduced and vice versa. This means that in general there is not one plan that Pareto dominates all other plans, but that there is a set of non-dominated plans.

In the context of clinical practice when a planner is presented the plans which are optimized by BRIGHT, the expert is able to pick one of the plans to use for the treatment based on the trade-off between target coverage and healthy tissue sparing. However, when the expert is not satisfied with the set of plans provided, the expert could change the protocol to have BRIGHT optimize plans based on this modified protocol. A possible reason for this could be when the expert concludes based on the presented plans that one of the organs is not spared sufficiently. In that case, the protocol will be modified. This does, however require BRIGHT to optimize the plans from scratch again.

The main goal of this research is to aim to decrease the time required after changing the protocol to find high-quality treatment plans. When the time to optimize after changing the protocol is decreased, it allows the expert to tweak the protocol to get plans which are more suited in a much faster and efficient way.

The effect of changing the protocol varies from patient to patient and from index to index. In some cases, a modification of the protocol does not affect the plans. The range of the aspiration value value for which this happens is called slack. When a aspiration value is modified within its slack value, there is no need to (re-)optimize for the modified protocol since the result will be similar. For modifications outside of the slack range, the effect is dependent on the anatomy of the patient and also dependent on which target is modified by what amount.

In order to lower the time required to converge to high-quality treatment plans a warm-start will be performed. A warm-start is performed by importing the population and the elitist archive of an initial run, after modifying the protocol. Even though these plans are optimized based on a different protocol, these plans still help BRIGHT in order to converge to high-quality treatment plans faster.

To further improve the time required to converge to high-quality treatment plans, the population and the elitist archive are modified before being imported. The goal of the modification is to modify the plans in such a way that they are more similar to the resulting plans of the second run. When plans are more similar to the final result, the expected result is a decrease in convergence time. One way this is attempted is based on the Euclidean distance to the organ of which the aspiration value is modified. Another way is based on a possible pattern between the sets of plans before and after modification. These attempts did not substantially improve the time required

because the changes were either too large, meaning that the non-modified plans outperformed the modified plans based on the modified protocol, or too small, meaning that there was no substantial difference between the non-modified and the modified plans.

The result of this research is that convergence time can be reduced greatly by using the population and the elitist archive of an initial run. The amount of time saved varies from patient to patient and from modification to modification, but in general it allows experts to feasibly modify the protocol without having to wait the full optimization time again. This research also shows that there might be more time to save by modifying the dwell times before importing the plans, however the approach used did not give a substantial decrease in convergence times.

Contents

1 Introduction

Nowadays, there are a lot of fields in which computers aid humans with processes, one of these fields is healthcare. As stated by Nazar et al., 2021, there are already a lot of artificial intelligence algorithms developed to be implemented in healthcare. One field within healthcare is to aid with planning radiation therapy to kill cancer cells while sparing surrounding healthy tissue as much as possible. There are various different types of radiotherapy, which vary based on the radiation method used or the location to which the radiation is applied. One specific radiation method, called brachytherapy, aims at delivering radiation from within the body. This is done by implanting a radiation source permanently or by temporarily implanting catheters (hollow needles) through which a radiation source can be guided. A treatment plan defines for what amount of time the radioactive source should be halted at which positions within the catheters, the longer the radiation is stopped at a point the more radiation is delivered to the surrounding tissue. This treatment plan aims to irradiate the tumor as much as possible while sparing the surrounding organs from receiving too much radiation. This means that within planning, there is a trade-off between covering the tumor and sparing the surrounding organs with regard to radiation. These plans are made by clinical experts. However, making a plan is a hard task, since in the case of high-dose rate brachytherapy for prostate cancer, there are a lot of parameters which have to be tweaked, for instance to distribute the radiation there are in general at least two hundred parameters which indicate how much radiation should be released at the given spots. A possible aid in this process is proposed by N. H. Luong et al., 2017, which uses the evolutionary algorithm MO-RV-GOMEA in order to generate a set of optimized plans ranging from good sparing/worse coverage to worse sparing/good coverage from which the clinical expert can choose the plan representing the preferred trade-off between coverage of the target volume and sparing of the organs at risk for the patient at hand. BRIGHT has been introduced in clinical practice and used since March 2020. The approach which is used is called inverse optimization. In inverse optimization an objective function is defined in order to determine the quality of a plan, this objective function is then either maximized or minimized (based on the definition of the objective function) in order to find high-quality plans.

This research aims at a request from clinical practice. When a set of plans is generated by MO-RV-GOMEA, in some cases, the clinical expert might want to change the protocol which is used in order to evaluate the quality of the plans (based on, for instance, observation of the first set of optimized plans). Currently, if the protocol is modified the optimization has to be restarted and will take the full optimization time again, which can take up to a few minutes. The goal of this research is to lower the time required to find good plans after modifying the protocol, such that an expert will get plans based on the modified protocol faster. A more detailed explanation of the aim of this research can be found in Chapter 2.

1.1. Background

1.1.1. Prostate HDR brachytherapy

In clinical practice, patients are treated for prostate cancer. One of the ways this is done, is by brachytherapy (Porter et al., 1995). The goal of brachytherapy is to guide one radioactive source near the cancer cells, to kill the cancer cells while at the same time minimizing radiation to the surrounding organs. The volume which has to be irradiated is called the planning target volume (PTV) which consists of the prostate (for this study the PTV equals the prostate, in some cases for an index PTV is used instead of prostate, this does not change the definition since both are the same volume), and the organs for which the radiation has to be minimized are called the organs at risk (OARs). Because the radiation source is guided into the body, brachytherapy is a form of internal radiation treatment. There are several types of brachytherapy for prostate cancer based on the way the radiation is applied, namely High-Dose-Rate (HDR), Low-Dose-Rate (LDR), and Pulse-Dose-Rate (PDR). This research will be focused on HDR, which makes use of a higher dose, but shorter time frame compared to LDR and PDR (Crook et al., 2020).

The main advantage of brachytherapy is that the source radiation is localized inside the body, near the tumor. This means that the radiation source can be placed close to the tumor, meaning that fewer of the surrounding organs will receive radiation. However, since other organs still surround the tumor, the radiation applied to the other organs should still be taken into account.

With HDR brachytherapy (HDR BT) the localized radiation is applied through hollow catheters. The catheters are inserted through the perineum and into the prostate, and into the seminal vesicles. Through these catheters, a radioactive source is guided and stopped at chosen locations for a pre-determined amount of time; these locations are called the dwell positions. The dwell positions are typically 2.5mm apart from each other. For the patients which are used in this research the amount of catheters varies between 17 and 24, and the amount of dwell positions depends on the number of catheters.

After the catheters are inserted, a scan (Computed Tomography (CT), Magnetic Resonance Imaging (MRI) or Ultrasonography) is made. Then an expert delineates (draws outlines of) the target volumes, the organs at risk and the catheters. After that, a clinical plan can be made. This is done by using the dwell positions and the delineated scan. A clinical plan consists of a set of dwell positions with their activation times. The longer the radioactive source stands still in a certain location, the more radiation is delivered at that location. The times that the radioactive source stands still are called dwell times. Based on this clinical plan, an afterloader will execute the plan by controlling the movements of the radiation source.

In Figure 1.1 the dwell positions and the dose calculation points for one patient can be seen from different angles. The radioactive source can be guided along the black dots and stand still at these points to deliver radiation to the surrounding area.

1.1.2. Clinical protocol

To assess the quality of a plan, a protocol is used. A protocol specifies a set of aspiration values that are preferred to be reached to have a good treatment plan. These aspiration values have an aim, in case of a sparing index the aim is to at least achieve the aspiration value and in case of a coverage index the aim is to score equal or lower than the aspiration value. A protocol consists of separate indices for target volumes and OARs and are either expressed in volume indices or dose indices. A volume index indicates how much volume of a given region of interest should receive either at least or at most (depending on whether it is aimed at coverage of the target volumes or sparing of the OARs) a certain percentage of the prescribed dose. A dose index is defined by the most irradiated sub-volume of an organ. The protocol is set up to spare the OAR and irradiate the tumor as much as possible.

As stated, the indices can be split up into dose indices and volume indices. Dose indices are defined as D_v^o, where v stands for the requested volume and o for the organ, and volume indices are defined as V_d^o, where d stands for the requested dose and o for the organ. Each index has an aspiration value. For coverage indices, the value of the index should be at least the aspiration value to satisfy the index, and for sparing indices, the value of the index should be at most the

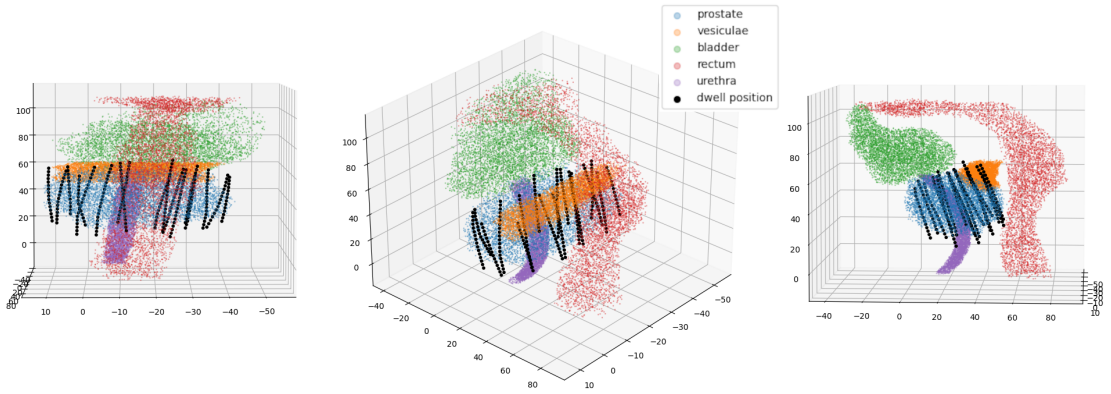


Figure 1.1: This figure shows three different angles of the sampled points of the regions of interest and the locations of the dwell positions for one patient. For visibility purposes each region of interest is visualized using 5.000 points which are randomly sampled from the region of interest.

value of the aspiration value to satisfy the index. These indices will be referred to as dose/volume indices (DVIs).

To clarify how DVIs work, two examples will be given. An example of a sparing volume index is the $V_{200\%}^{prostate}$ <15%. For a plan to satisfy this volume index, the volume with a planned dose of 200% of the prescribed dose (which is 15Gy * 200% = 30Gy) should be smaller than 15% of the total volume of the prostate.

An example of a coverage dose index is the $D_{90.0\%}^{prostate}$ >15Gy. For this coverage dose aim to be satisfied, at least 90% of the prostate should have a planned radiation dose of at least 15Gy.

In some cases, a planner does not fully adhere to the protocol because the clinical protocol is not working well enough for the specific patient at hand. In some cases, the implants are not inserted correctly or do not reach deep enough, making it hard or impossible to fulfill the aim for the coverage of the seminal vesicles.

The protocol used in this research is the protocol that is currently being used by the Amsterdam University Medical Centers (AUMC) (and is similar to the protocol described by Henry et al., 2022). This protocol will be referred to as the AUMC protocol and will be the only protocol that is used in this research. The AUMC protocol can be seen in Table 1.1.

1.2. Evolutionary Algorithms

In order to deal with the HDR BT planning problem, an algorithm needs to be used. For this research, the problem is solved using an evolutionary algorithm (EA). EAs are a population-based search heuristic. One of the main reasons to tackle the HDR BT planning problem with EAs, is that the HDR brachytherapy problem is a multi-objective problem and EAs are state-of-the-art in multi-objective optimization. Since the HDR BT planning has a clear trade-off in objectives multi-objective optimization helps tackle this problem. The trade-off in the HDR BT planning is that the target volume should be irradiated sufficiently while still sparing the surrounding organs (organs at risk). The reason to use the multi-objective definition of the HDR BT planning is that within clinical practice also multiple objectives are used, based on coverage and sparing, to assess the quality of a plan.

The concept of EAs is inspired by biological evolution. An example of an EA is a classic genetic algorithm (GA) (Holland, 1992. A classic GA starts with a population of individuals. Each individual has a fitness value (which is determined by a function that determines the quality of the individual). Using existing individuals within the population, new solutions are generated by variation. Variation is used to explore the solution space and generates new individuals using the parent individuals and could also use randomness to create/generate solutions. There are several ways in which variation can be applied. Examples of this are cross-over (in which values

AUMC protocol				
Prescribed dose: 15Gy				
Coverage indices				
Region of interest	Type	Requested	Target	Definition
prostate (or PTV)	volume	100%	>95%	$V_{100\%}^{prostate}$ >95%
prostate (or PTV)	dose	90%	>15Gy	$D_{90\%}^{prostate}$ >15Gy
vesiculae	volume	73%	>95%	$V_{73\%}^{vesiculae}$ >95%
Sparing indices				
Region of interest	Type	Requested	Target	Definition
prostate (or PTV)	volume	150%	<40%	$V_{150\%}^{prostate}$ <40%
prostate (or PTV)	volume	200%	<15%	$V_{200\%}^{prostate}$ <15%
bladder	dose	1cm^3	<13Gy	$D_{1cm^3}^{bladder}$ <13Gy
bladder	dose	2cm^3	<12Gy	$D_{2cm^3}^{bladder}$ <12Gy
rectum	dose	1cm^3	<11Gy	$D_{1cm^3}^{rectum}$ <11Gy
rectum	dose	2cm^3	<9.5Gy	$D_{2cm^3}^{rectum}$ <9.5Gy
urethra	dose	30%	<16.5Gy	$D_{30\%}^{urethra}$ <16.5Gy
urethra	dose	0.1cm^3	<18Gy	$D_{0.1cm^3}^{urethra}$ <18Gy

Table 1.1: AUMC clinical protocol. When a requested dose it is a percentage based on the prescribed dose (i.e., the requested dose for $V_{100\%}^{prostate}$ is 100% of 15Gy, therefore it is 15Gy)

of the parent are partially copied over) or mutation (in which random changes are applied). After this process, the selection part of the algorithm is executed. In the selection, a set of individuals is selected, which will be kept, while the other individuals are discarded. The most common way of selection is based on fitness values, meaning that the individuals with better fitness scores are selected.

The process of variation and selection is called a generation. These generations will be executed until termination. The algorithm will terminate based on a termination condition, which can be, for instance, a value to reach (a target fitness value), a set number of generations, or a set amount of time.

In general, there are four steps in every evolutionary algorithm. These steps are the initialization, the selection (elimination), the variation, and the termination. At first, a population has to exist in order to start the evolutionary algorithm. The initial population can, for instance, be randomly generated. These individuals which are in the initial population will also be evaluated with the fitness function. After the initialization, the selection will be performed. Within the selection, the best solutions of the population are selected. Based on these individuals, variation is applied.

The goal of variation is to explore and find better solutions, which is often done by a combination of cross-over and mutations. The goal of cross-over is to take parts of existing individuals and use them to construct higher-quality offspring solutions, so it basically copies a part of the original individual into the new individual. In order to make sure there is sufficient exploration, mutation is sometimes used. A mutation is a change that is not based on an existing individual but a way to explore a solution that would not be found by combining existing individuals. After the variation selection happens again, these two steps will be executed until the evolutionary algorithm is stopped. Each step of selection and variation is called a generation. There are several conditions under which an evolutionary algorithm can be terminated. These conditions are called termination conditions. Termination conditions could be, for instance, run time, number of generations, or when a new population does not improve on the previous population.

One of the main reasons to use evolutionary algorithms is their ability to use multiple objectives while optimizing. The use of multiple objectives is called multi-objective optimization (MO) (Deb and Deb, 2014). EAs in general are useful for solving multi objective optimization problems, since the population can be used to track multiple solutions. The process when optimizing on multiple objectives is similar to optimizing a single objective. However, when selection is performed, the evolutionary algorithm will select in a different way. One possible way is Pareto domination. When Pareto dominance is applied to two individuals which have multiple objectives (and therefore multiple fitness scores), an individual Pareto dominates another individual if and only if it is at least equal for all objectives than the other individual and strictly better in at least one of the objectives. This Pareto dominance can be used in the selection of an evolutionary algorithm in order to compare individuals. Given a set of individuals, the Pareto set is defined to be a set in which all individuals are non-dominated by any other individual, so when an individual is not in the Pareto set, it means that this individual is dominated by at least one other individual.

1.2.1. MO-RV-GOMEA

The Gene-pool Optimal Mixing Evolutionary Algorithm (GOMEA) is an evolutionary algorithm for solving discrete optimization problems (Thierens and Bosman, 2011). The version of GOMEA which is used in this research is the multi-objective (MO) (N. H. Luong et al., 2014) and real-valued (RV) (Bouter, Alderliesten, Witteveen, et al., 2017) variant, named MO-RV-GOMEA (Bouter, Luong, et al., 2017). One of the strengths of GOMEA is that it can exploit relations between the input parameters. These relations are used in the variation operator, which is named Gene-pool Optimal Mixing (GOM). Another property of GOMEA is that it divides the population into clusters which are used to make sure that solutions from different parts of the front are not mixed, since these solutions are likely to have different characteristics.

Within GOMEA, a linkage tree is used in order to exploit the relation between the input parameters. The linkage tree is a tree-based structure and is a way to hierarchically cluster input parameters based on the relation with other input parameters. Each node consists of a set of input parameters called a linkage set. For the children of a linkage set, it holds that their union is equal to the linkage set of their parent. A linkage set represents a set of related input parameters. The linkage tree can be pre-defined or calculated based on the population. This linkage tree is used in the variation step of GOMEA to be able to sample solutions based on changing input parameters in individuals which are related. The nodes of the linkage tree are linkage sets.

For variation in GOMEA, the τ% (for this research always 35) best solutions are divided into clusters. For each of these clusters, a Gaussian distribution is estimated for each linkage set. A linkage set contains a subset of the input variables, and the variables within a linkage set are related to each other. Based on the generated distributions (estimated based on the aforementioned clusters), new partial solutions are sampled (a partial solution is a modification of a parent solution, only changing the variables which are defined in the partial solution), and these new solutions have changes based on linkage sets. GOMEA uses Gene-pool Optimal Mixing (GOM) (Thierens and Bosman, 2011), which samples partial solutions based on linkage sets and only accepts a change if it improves the fitness.

Because of the way GOMEA applies partial solution variation for some problems, it allows

the use of partial evaluations (Bouter, Luong, et al., 2017). In the variation step of GOMEA, partial solutions (containing values for only a set of input parameters, not all) are applied. When evaluating whether the partial solution improves the current individual, only a partial evaluation can be executed. This can be used to be able to accept or reject a partial solution quicker than with full evaluations, which is, therefore, more efficient.

GOMEA makes use of an elitist archive to keep track of the best solutions which are known at that moment. The best solutions are defined to be the solutions that are non-dominated. For an individual to be non-dominated, it has to hold that there is no other known individual that performs better for at least one objective and at least equal for all other objectives.

1.3. BRIGHT

The application of MO-RV-GOMEA on HDR BT planning is called BRachytherapy via artificially Intelligent GOMEA Heuristic based Treatment planning (BRIGHT). BRIGHT is used for finding high-quality treatment plans varying from high coverage/low sparing to low coverage/high sparing (N. H. Luong et al., 2017). The application of MO-RV-GOMEA on HDR BT planning has also been proven to give good feasible plans, in most cases, by an observer study (a research conducted in which experts assesed/compared plans), equal to or better than the clinical plans previously made using the available software tools at the time (Maree et al., 2019).

Besides allowing multi-objective optimization (Bouter, Luong, et al., 2017), MO-RV-GOMEA also allows real-valued inputs. The multi-objective optimization is an advantage since, within the brachytherapy planning problem, there is a clear trade-off, namely between covering the planning target volume and sparing the organs at risk. Even though it would be possible to combine coverage and sparing within one objective, the advantage of the trade-off which is presented in the output would be lost. When using two objectives, one for coverage and one for sparing, the evolutionary algorithm can provide a set of plans, ranging from low coverage and high sparing to high coverage and low sparing. This gives an expert a set of plans to choose from. This way, the expert can decide on the trade-off between coverage and sparing instead of the algorithm providing only one solution which is based on pre-determined weights which weigh the importance of coverage and sparing. Having a set of plans which trade-off sparing and coverage is important since based on the patient at hand an expert might decide to pick a different trade-off. Another advantage of the multi-objective optimization approach is that it gives a good insight on what kind of plans can be obtained and what the trade-offs between those plans are and what could change of a plan without being worse (dominated) than an other plan.

In BRIGHT, for MO-RV-GOMEA, one plan is one individual. Each individual consists of one real-valued parameter for each dwell position. These values are called dwell times. The population is a set of individuals which will be used to explore the search space for possible treatment plans. When referring to the population of BRIGHT P will be used.

The initial population of BRIGHT is randomly initialized. For each plan, each dwell time will get a randomly generated value that is uniformly distributed between 0 and 1.

To be able to determine the quality of a plan, the values of the indices have to be calculated. For that, it is required to know how much radiation dose is delivered to each of the regions of interest. To calculate the planned radiation to an organ, a set of dose calculation points (DCPs) is sampled from each of the organs. These DCPs are randomly sampled within the organs based on the delineations made by the expert (based on the initial CT or MRI scan). For a full evaluation of an individual, the radiation in each dose point will be calculated. Based on the DCPs and dwell positions, a dose-rate matrix is created, which is used to calculate the dose distribution. To find the planned dose to the DCPs, the dose-rate matrix is multiplied by the dwell time vector, which contains all the parameters of an individual.

In order to assess the fitness of an individual (plan), the set of dwell times will be evaluated according to the protocol. Each of the indices in the protocol will be evaluated and will receive a score, this score is then normalized with regards to their relative performance, and the indices are then grouped into two groups, namely sparing indices and coverage indices. Of both groups, the worst-performing index is chosen, and those make up the objective values for this individual. The

objective determined by the worst-performing sparing index is called Least Sparing Index (LSI), and the objective determined by the worst-performing coverage index is called Least Coverage Index (LCI).

In order to be able to determine the LCI value and the LSI value, the values of the indices have to be compared. Because of this, normalization of the values is required. The normalization is done in order to normalize the distance between the actual value of the index with regard to the aspiration value. Without normalization the indices are harder to compare, since their maximum performance (scores) do differ and the aspiration values do differ, for instance a 1% violation of an aspiration value of 15% of an organ is more significant than a 1% violation of an aspiration value of 95%. The normalized values can be calculated using the functions $\delta_c(V_d^o)$ for coverage indices and $\delta_s(D_v^o)$ for sparing indices. V_d^o stands for a volume index based on organ o with requested dose d. D_v^o stands for a dose index based on organ o with requested volume v. To indicate the aspiration value for an index (as defined in Section 1.1.2) $V_d^{o,target}$ and $D_v^{o,target}$ will be used (so for instance $V_d^{o,target}$ for $V_{150\%}^{prostate}$ is 40% of the volume of the prostate). To indicate the value or score of a DVI for a specific plan x, $V_d^{o,x}$ and $D_v^{o,x}$ will be used. Because volume indices can be defined in either a percentage of the total volume of the organ or as an absolute volume, when normalizing a volume index this is done with regards to the total volume of the organ. This volume will be indicated as Vol_o where o is the organ. The definitions of these functions can be seen below.

$$\delta_c(D_v^{o,x}) = \frac{-(D_v^{o,target} - D_v^{o,x})}{0.3(D_v^{o,target})} \tag{1.1}$$

$$\delta_c(V_d^{o,x}) = \begin{cases} \frac{-(V_d^{o,target} - V_d^{o,x})}{100 - V_d^{o,target}}, & \text{if volume unit of } V_d^o \text{ is \%} \\ \\ \frac{-(V_d^{o,target} - V_d^{o,x})}{Vol_o - V_d^{o,target}}, & \text{elseif volume unit of } V_d^o \text{ is cm}^3 \end{cases} \tag{1.2}$$

$$\delta_s(D_v^{o,x}) = \frac{D_v^{o,target} - D_v^{o,x}}{D_v^{o,target}} \tag{1.3}$$

$$\delta_s(V_d^{o,x}) = \frac{V_d^{o,target} - V_d^{o,x}}{V_d^{o,target}} \tag{1.4}$$

Using the normalized values, the LCI and the LSI of a plan x can be calculated in the following way

$$LCI(x) = min\{\delta_c(V_{100\%}^{prostate,x}), \delta_c(D_{90\%}^{prostate,x}), \delta_c(V_{73\%}^{vesiculae,x})\} \tag{1.5}$$

$$LSI(x) = min\{\delta_s(V_{150\%}^{prostate,x}), \delta_s(V_{200\%}^{prostate,x}), \delta_s(D_{1cm^3}^{bladder,x}), \delta_s(D_{2cm^3}^{bladder,x}), \delta_s(D_{1cm^3}^{rectum,x}), \delta_s(D_{2cm^3}^{rectum,x}), \delta_s(D_{30\%}^{urethra,x}), \delta_s(D_{0.1cm^3}^{urethra,x})\} \tag{1.6}$$

In order to ensure BRIGHT explores the solution space in which useful plans are found, it makes use of a constraint mechanism. Based on the scores of the indices, BRIGHT calculates a constraint value. If the constraint value is 0, then the plan does satisfy the constraints. The higher the constraint value, the further the individual is from being within the constraints. Constraint domination is used when plans are being compared, i.e., a plan with a constraint value non-equal to zero is always dominated by a plan with a constraint value of zero. If two individuals have a constraint value non equal to zero, the one with the lowest constraint value dominates the other individual. The constraint value is calculated by the sum of the violation of hard constraints (of indices of the protocol) and the violation of the LCI and LSI lower bounds. The AUMC protocol (which is used in this research) does not have any hard constraints, so that part is always zero. The second part of the constraint value is whether the individual violates the lower bounds of

LCI and LSI (which are set to -0.2 for both however this is for the non-normalized LCI and LSI values), the reason for this constraint is that plans outside of this constraint range are so far from the golden corner that these plans are not interesting for the experts because they either irradiate too much or too little.

In the elitist archive, the plans which are non-dominated by any other dominated plan are stored. For BRIGHT, this means that an individual is non-dominated if there is no plan with a better LCI and the same LSI (or better) and if there is no plan with a better LSI and the same LCI (or better). If a new individual enters the elitist archive, all plans that are dominated by that new plan are removed from the elitist archive. When referring to the elitist archive of BRIGHT $\mathcal{E}$ will be used. Theoretically the elitist archive can grow very large, since there are a lot of possible solutions. In order to prevent the elitist archive from growing too large objective-space discretization is used, as defined by H. N. Luong and Bosman, 2012. When adding a new solution to the elitist archive, it should not only be non-dominated by any other solution in the archive, but also not be too similar with regards to objectives (as described in H. N. Luong and Bosman, 2012, this is done by placing a solution in a hypercube within the objective space).

As explained in Section 1.2.1, MO-RV-GOMEA exploits linkage information by using a linkage tree. In general, the linkage tree can be learned. However since a measure for the dependence between the dwell positions (the variables for MO-RV-GOMEA) is known based on problem-specific knowledge, the linkage tree can be calculated at the start. This can be done by creating the linkage sets based on the Euclidean distance between the dwell positions (dwell positions that are close to each other are more likely to be in the same linkage set). Since, in general, patients treated in clinical practice of the AUMC have over 200 dwell positions, the minimum linkage set size is five variables.

At the variation step based on FOS elements, new partial solutions are sampled and then applied to existing solutions. Because only a part of the dwell times are changed, a lot of the calculation of the doses in the organs can be skipped. For a full evaluation of a plan, the dose applied by each dwell position to each dose calculation point is calculated. This is done by multiplying the dose-rate matrix (which defines the effect of a dwell position on a dose calculation point) with the dwell times of a plan. When only a part of the dwell times changes, only the dose applied to those dwell positions has to be re-calculated for the evaluation, which saves time, such an evaluation where only a part of the dose distribution has to be re-calculated is called a partial evaluation.

There are several ways in which BRIGHT can be terminated. The one used in this research will be number of generations. A good termination condition is important, since if the algorithm terminates late then time is wasted and if the algorithm terminates early the plans are not as good as they could be. In general, at the start plans do have a lot of improvement when comparing with a previous generation, however, this improvement between the generations becomes smaller over time. This continues until plans barely change or improve over time, then the set of plans (and therefore the algorithm) is then considered to be converged. The amount of generations for the termination condition is picked in such a way that convergence can be assumed.

1.3.1. Exponential weighting

BRIGHT uses exponential weighting in order to also optimize for other indices which are not the worst (Bouter et al., 2019). This is done because if only the worst-performing indices are being optimized, the other indices will not be optimized. The reason that not every individual index is optimized on when not using exponential weighting, is that some indices never become the worst performing sparing or coverage index and optimization is based on LCI/LSI.

Especially when looking at indices that are easier to satisfy than other indices, these indices will quickly have a better score than other indices, meaning optimizing them does not benefit the objectives. Therefore these indices will have lower scores than what they potentially could be.

To allow optimization on indices that are not the worst-performing indices BRIGHT uses exponential weighting. With this, the LCI and LSI are not just based on the worst-performing indices, but with an exponential factor, other indices are also included.

With exponential weighting, every index is multiplied by a factor and added. The best performing index within a category (sparing or coverage) gets a factor of 1. The second-best performing index gets a factor of 10, and the third-best performing index gets a factor of 100, etc. This means that every index gets ten times the weight of the previous best index.

In the following equation $w_c(\delta_c(V|D^o_{d|v}))$ is defined, this is a function to get the weighted variant of $\delta_c(V|D^o_{d|v})$, which is a normalized dose or coverage index ($\delta_c(V^o_d)$ or $\delta_c(D^o_v)$). Using this function the weighted LCI, LCI_w, can be found for a plan x.

$$LCI^{\text{sorted}}(x) = sort\{\delta_c(V^{prostate,x}_{100\%}), \delta_c(D^{prostate,x}_{90\%}), \delta_c(V^{vesiculae,x}_{73\%})\}$$

$$W_c = [\lambda^0, \lambda^1, \lambda^2]$$

$$W_{c,total} = \sum_{n=0}^{2} \lambda^n$$

$$w_c(\delta_c(V|D^{o,x}_{d|v})) = \frac{\delta_c(V|D^{o,x}_{d|v}) * W_c[i]}{W_{c,total}} \text{ where } i \text{ is the index of } \delta_c(V|D^{o,x}_{d|v}) \text{ in } LCI^{\text{sorted}}$$

$$LCI_w(x) = \sum_{n=0}^{3} w_c(LCI^{\text{sorted}}(x)[n])$$

LSI_w for a plan x can be calculated in a similar way.

$$LSI^{\text{sorted}}(x) = sorted\{\delta_s(V^{prostate,x}_{150\%}), \delta_s(V^{prostate,x}_{200\%}), \delta_s(D^{bladder,x}_{1cm^3}), \delta_s(D^{bladder,x}_{2cm^3}), \delta_s(D^{rectum,x}_{1cm^3}), \delta_s(D^{rectum,x}_{2cm^3}),$$
$$\delta_s(D^{urethra,x}_{30\%}), \delta_s(D^{urethra,x}_{0.1cm^3})\}$$

$$W_s = [\lambda^0, \lambda^1, \lambda^2, \lambda^3, \lambda^4, \lambda^5, \lambda^6, \lambda^7]$$

$$W_{s,total} = \sum_{n=0}^{7} \lambda^n$$

$$w_s(\delta_s(V|D^{o,x}_{d|v})) = \frac{\delta_s(V|D^{o,x}_{d|v}) * W_s[i]}{W_{c,total}} \text{ where } i \text{ is the index of } \delta_s(V|D^{o,x}_{d|v}) \text{ in } LSI^{\text{sorted}}$$

$$LSI_w(x) = \sum_{n=0}^{8} w_c(LSI^{\text{sorted}}(x)[n])$$

Picking different values for λ will result in different ways of evaluation. With $\lambda = 1$ all indices would contribute the same amount to LCI or LSI. When λ grows, the more important the worst performing index becomes. For this research, the LCI_w and the LSI_w are used as the objective, so if LCI or LCI is stated, it is the weighted variant LCI_w or LSI_w. Based on Bouter et al., 2019, for λ a value of 10 is chosen.

2

Problem statement

This research is inspired by a use case from clinical practice. When BRIGHT is used in clinical practice it optimizes a set of plans which are then presented to the expert. In some cases the expert might not be satisfied with the set of plans which is presented. In that case currently, the expert could pick one plan and modify it by hand (which takes time and will likely decrease the quality of the plan) or modify the protocol and restart the optimization from scratch (which takes the full optimization time again). The goal of this research is to improve this process and to allow experts to modify the protocol after a run without having to optimize BRIGHT from scratch again, this allows an expert to be able to tweak the protocol based on the results and to get new plans based on the modified protocol much faster than optimization from scratch.

So, the aim of this research is to reduce the time until convergence in case the protocol changes at a given moment. Currently, when using MO-RV-GOMEA on the brachytherapy planning problem, when the protocol changes (and therefore the objective space changes), the optimization has to be executed from scratch again. The goal is to improve on that, to be able to adapt dynamically and make sure the convergence is as fast as possible. Besides that, it is also important that similar results will be obtained when trying to converge more quickly compared to when running from scratch with the adapted protocol.

The protocol can be changed such that it is either harder to satisfy all indices or sucht that it is easier to satisfy all indices.

In some cases, the protocol might be changed such that is is easier to be fullfilled. For example, when the catheters do not reach deep enough into the seminal vesicles, there are no dwell positions available close to the seminal vesicles, making it too hard to satisfy the coverage index $V_{73\%}^{vesiculae}$ while still also satisfying the coverage indices. In that case, the aspiration value for this index can be decreased, making it easier to satisfy.

To be able to tackle this problem, first, the scope of the changes in the objective space has to be defined. The objective is defined as the worst-performing indices for both sparing and coverage, so in most cases, the objective values are mostly defined by two of the indices. This means that for the other indices, there is a margin in which the index can be made harder to achieve without (or barely) influencing the objective values. Since changing these objectives within this margin does not influence the problem, these changes will not be considered. The changes should have a substantial impact on the objectives. However, in clinical practice, it is also important to know when a change is not substantial enough to have impact. Therefore, a way to detect whether a change requires optimization again or whether there are no changes would benefit the clinical practice.

3

Influence of changing the protocol

In order to speed up the time required to converge to a set of high-quality treatment plans after changing the protocol, first, the effect of different changes on the protocol needs to be evaluated. Evaluating the influence of changing the protocol will give more insights into the effect of a protocol change. This knowledge can then be used in order to speed up convergence after changing the protocol.

3.1. Method

Since the objective space is based on LSI and LCI, which are directly related to the protocol, a change in the protocol has an impact on the objective space. This is because changing the protocol has an influence on the distance of the indices to their respective aspiration value as defined in the protocol. So a change in the protocol will have an effect on the indices which are used to calculate LCI and LSI.

In general, if a coverage index is made more difficult to achieve by increasing the aspiration value value, the index will score relatively worse compared to the other unchanged coverage indices. This means that after changing the aspiration value value, this index will have more influence on the LCI after the change.

The same holds for sparing indices. If a sparing index is made harder by decreasing the aspiration value value, the index will perform worse and be more influential in determining LSI.

It is, however, not always the case that by increasing the difficulty of an index, the LCI or LSI will change. This could happen since only the worst-performing indices will determine the LCI and LSI. This margin, in which an aspiration value can change without substantially influencing LCI and LSI, will be called the slack on an index. The concept of slack is defined and explained in Section 3.2.

In order to check the influence of a change in the protocol, two runs will be compared, one of which is a run from scratch on the original protocol. Then, another run from scratch is executed but with the change in the protocol. This means that both runs will have their initial plans randomly initialized. This can be seen in Figure 3.1. The notation is explained in the next section, Section 3.1.1.

3.1.1. Notation

The protocol denoted ρ, which is used in this research, is always (based on) the AUMC protocol, which is defined in Section 1.1.2. When the protocol is modified it will be denoted as ρ', indicating a modified version of the AUMC protocol. The modification is defined when ρ' is used.

As stated in Section 1.3, the notation used for the population and elitist archive are P and $\mathcal{E}$, respectively. As explained in the previous paragraph, the protocol which is used can be the standard protocol or a modified version of it. If the population and archive are based on the original

protocol, this will be indicated as P^{ρ} and $\mathcal{E}^{\rho}$. On the other hand, if the protocol is modified, it will be indicated as $P^{\rho'}$ and $\mathcal{E}^{\rho'}$ for the population and the elitist archive, respectively.

When BRIGHT is initialized, it usually generates a random population, but this population is not always randomly generated, especially in this research. In the case that the initial population, indicated by P_{in},of the first run is randomized this will be indicated as $P_{in} \leftarrow$ randomPopulation (the way in which the random initialization is executed can be seen in Section 1.3).

The initial elitist archive in BRIGHT, denoted E_{in}, is generally empty, and will be filled throughout optimization. An empty initial elitist archive is denoted $E_{in} \leftarrow \emptyset$.

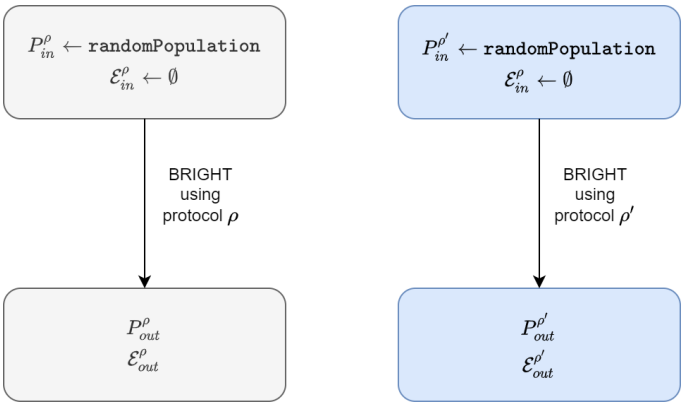


Figure 3.1: Structure of execution for determining the influence of a change in the protocol on the results of BRIGHT. The notation is defined in Section 3.1.1

3.2. Slack

For this section, the definitions of LCI and LSI in their unweighted forms are considered (so not the LCI_w and LSI_w as explained in Section 1.3.1). So for slack calculation, LCI and LSI will be considered to be only the value of the worst-performing indices. This means that there could be slight changes within LCI and LSI, even when the change of the aspiration value of the index is within the slack. The application of exponential weighting on the concept of slack will be explained in section 3.2.1.

The objectives (LCI and LSI) are determined by the worst-performing indices. Because of that, any indices which are not the worst in their category (sparing or coverage) before and after the change in protocol will not have an influence on the LCI and LSI value.

This margin will be defined as slack. The slack of a single plan can be calculated by checking which indices are the worst-performing indices and therefore make up the LCI and LSI values. These indices do not have any slack since if these indices are made more difficult to achieve, the LCI or the LSI value will be changed. The notation defined in Section 3.1.1 will be used to help define the slack.

To calculate the slack, the normalized value of the given index should be set equal to the normalized value of the worst index. The slack is defined as $s_{D_v^{o,x}}$ for a dose index and as $s_{V_d^{o,x}}$ for a volume index for a plan x. The slack is the distance of the score of an index to the score of the worst-performing index. The slack can be calculated for each plan x by solving the following equations for $s_{D_v^{o,x}}$, or $s_{V_d^{o,x}}$, respectively.

$$LCI(x) = \frac{(D_v^{o,target} + s_{D_v^{o,x}}) - D_v^{o,x}}{0.3(D_v^{o,target} + s_{D_v^{o,x}})} \tag{3.1}$$

$$LCI(x) = \begin{cases} \frac{(V_d^{o,target} + s_{V_d^{o,x}}) - V_d^{o,x}}{100 - (V_d^{o,target} + s_{V_d^{o,x}})}, & \text{if volume unit of } V_d^o \text{ is \%} \\ \frac{(V_d^{o,target} + s_{V_d^{o,x}}) - V_d^{o,x}}{Vol_o - (V_d^{o,target} + s_{V_d^{o,x}})}, & \text{elseif volume unit of } V_d^o \text{ is cm}^3 \end{cases} \tag{3.2}$$

$$LSI(x) = \frac{-((D_v^{o,target} + s_{D_v^{o,x}}) - D_v^o)}{(D_v^{o,target} + s_{D_v^{o,x}})} \tag{3.3}$$

$$LSI(x) = \frac{-((V_d^{o,target} + s_{V_d^{o,x}}) - V_d^o)}{(V_d^{o,target} + s_{V_d^{o,x}})} \tag{3.4}$$

For the worst-performing DVIs it is defined that the slack ($s_{D_v^o}$ or $s_{V_d^o}$) is equal to zero since these DVIs are already the worst-performing DVIs, so the distance to themselves is 0.

If the equations are solved for $s_{D_v^o}$ or $s_{V_d^o}$ then the slack of an index within a plan can be calculated by plugging the values into the formula.

$$s_{D_v^{o,x}} = \frac{(10 - (3 * LCI(x)) * D_v^{o,target} - 10 * D_v^{o,x}}{3 * LCI - 10} \tag{3.5}$$

$$s_{V_d^{o,x}} = \begin{cases} \frac{-V_d^{o,target}(LCI(x)+1) + V_d^{o,x} + 100 * LCI(x)}{LCI(x)+1}, & \text{if volume unit of } V_d^o \text{ is \%} \\ \frac{-V_d^{o,target}(LCI(x)+1) + V_d^{o,x} + Vol_o * LCI(x)}{LCI(x)+1}, & \text{elseif volume unit of } V_d^o \text{ is cm}^3 \end{cases} \tag{3.6}$$

$$s_{D_v^{o,x}} = \frac{D_v^{o,x}}{LSI(x)+1} - D_v^{o,target} \tag{3.7}$$

$$s_{V_d^{o,x}} = \frac{V_d^{o,x}}{LSI(x)+1} - V_d^{o,target} \tag{3.8}$$

One of the main advantages of this method of the calculation of the slack is that the run time is not of substantial length when comparing it to the run time of BRIGHT. It also scales linearly in run time with regard to the number of plans. The slack can be calculated based on a set of plans and can then be used in clinical practice, for instance, to save time when a change to the protocol within the slack is made (since, in that case, BRIGHT does not have to run since the results will be similar). Besides deciding whether time can be saved by not re-running BRIGHT, the slack values can also be used to see which DVIs are the hardest or easiest to achieve and analyze the implications of using different protocols on different patients.

3.2.1. Exponential weighting with slack

The slack described in Section 3.2 is without taking the exponential weighting into account. Without the exponential weighting, the LCI and LSI would not change when an aspiration value is changed within its slack. However, with the exponential weighting, every index contributes some (possibly negligible) value to the LCI or LSI.

However, with exponential weighting, the worst-performing index is still the most determining factor. There are 3 coverage indices, meaning a total weight of $1 + 10 + 100 = 111$, the worst-performing index gets assigned a weight of 100 and therefore determines $\frac{100}{111} * 100\% = 90.09\%$ of the LCI. The protocol has eight sparing indices, so the total sum of weights is 11111111. The

worst-performing sparing index has a weight of 10000000, so that means that the worst-performing sparing index determines $\frac{10000000}{11111111} * 100\% = 90.00\%$ of the LSI.

The slack is defined as the margin in which an aspiration value can be adjusted while still not being the worst performing index. This means that if an aspiration value is changed within its slack that it would be at worst the second-worst performing index, meaning that it will have approximately 9% influence on the LCI or LSI. However, this does not mean the LCI or LSI could change with 9% of itself since the index will have a 9% influence on the LCI or LSI. This means that when taking into account exponential weighting when calculating the slack, the LCI or LSI value can change with at most 9% of the change in the score of the index, which is based on the aspiration value change.

4

Effects of importing individuals

There are several ways in which the optimization, following a change in the clinical protocol, can be improved. One possible way to improve the optimization is to make use of information from the initial run. Even though this run is based on a different protocol, there is still a lot of valuable information which can be used. Importing individuals is a way to give the algorithm a so-called 'warm-start', meaning that optimization will be aided by the imported individuals at the start. This is the opposite of a 'cold-start' in which the individuals are randomly initialized.

In a sense, this way, the optimization is continued since the individuals of the initial run will be imported at the start of the new optimization. This is mainly useful when the changes in the protocol are small because, in that case, the solutions will already have relatively good fitness values compared to the default random initialization.

There are two different sets of plans which will be imported, namely the population and the elitist archive. As stated before, the population will replace the randomly initialized population. The elitist archive is generally initially filled with the Pareto set of the population. When importing the elitist archive, the elitist archive will contain the Pareto set of the population and the imported elitist archive combined.

4.1. Importing the population

The goal of the population is to explore the solution space and find solutions that improve on the currently known solutions. The initialization of the population is done randomly ($P_{in} \leftarrow$ randomPopulation). After each generation, the population will be either as good as the previous generation or better since only individuals (plans) which are improved will be accepted (this is one of the properties of MO-RV-GOMEA).

In order to improve the convergence speed of a second run which is executed on the modified AUMC protocol (ρ'), the converged population of the first run, which is based on the AUMC protocol, can be used ($P_{in}^{\rho'} \leftarrow P_{out}^{\rho}$). Therefore, instead of the usual random initialization of the dwell times, the initial population will be imported from an already converged run.

The population consists of a set of plans. Each plan is defined by a list of dwell times (one for each dwell position), which are imported. The plans also have their objective values and the scores for the indices stored. However, these are not imported, as they are no longer valid due to the change in protocol.

As stated before, the imported population P_{out}^{ρ} will replace the randomly initialized population. This also means that the elitist archive will be built based on the imported population. Therefore, the initial elitist archive will consist of the non-dominated individuals from the population rather than the randomly initialized individuals.

The concept of importing the population is displayed in Figure 4.1.

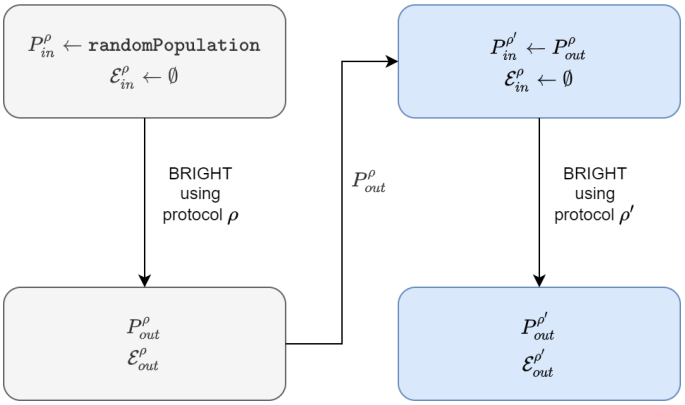


Figure 4.1: Structure of execution for importing the population in order to improve convergence speed of the second run. The notation is defined in Section 3.1.1

4.2. Importing the elitist archive

The elitist archive is a part of the algorithm, which consists of the best-known solutions since for a plan to be in the elitist archive, it has to hold that there is no other known plan which Pareto dominates that plan. During generation, for each cluster, a part of the solutions from the elitist archive is copied into the population, and based on the updated population the clustering and the estimation of the distributions is done. Therefore the elitist archive also influences the variation process, which makes use of the estimation of the distributions.

After a run of BRIGHT, the elitist archive consists of all the best solutions which are found during the optimization, i.e., all non-dominated individuals. In general, after convergence, the elitist archive has more individuals than the population and therefore has more information. Since the elitist archive of an initial run on the original AUMC protocol ($\mathcal{E}^{\rho}_{out}$) consists of the best-known solutions, it might be useful to use as input for a second run in which the protocol is modified (ρ').

After importing the elitist archive ($\mathcal{E}^{\rho'}_{in} \leftarrow \mathcal{E}^{\rho}_{out}$), first the imported elitist archive will be evaluated with regards to the modified protocol (ρ'). After the evaluation of the elitist archive, not all individuals of the elitist archive will be non-dominated any longer. This occurs because the change in the protocol will not have the same effect on every individual, meaning that some individuals might decrease more or increase less with regard to their LCI/LSI scores. If this happens, it could be that those individuals are now dominated by other individuals from the elitist archive, meaning that those dominated individuals are discarded.

Another way in which the elitist archive size shrinks compared to the result of the initial run to what is left of the elitist archive after initialization is the individuals, which will have a constraint value greater than 0. This means that the plan is infeasible because it has a non-normalized LCI or LSI value smaller than -0.2. These plans with a constraint value greater than 0 did have a constraint value of 0 in the output of the initial run ($\mathcal{E}^{\rho}_{out}$), however, since the protocol changed, the LCI and LSI values change, and since plans with a constraint value greater than 0 are dominated by plans with a constraint value which have a constraint value of 0, these plans will be discarded from the archive.

Before the execution of the first generation, the non-dominated individuals (non-dominated by both the elitist archive and the population) of the population are added to the elitist archive. When importing the elitist archive in a second run on a modified protocol ($\mathcal{E}^{\rho'}_{in} \leftarrow \mathcal{E}^{\rho}_{out}$), when the population is randomly initialized on the second run ($P^{\rho'}_{in} \leftarrow$ randomPopulation) none of the individuals of the population will generally enter the elitist archive since randomly initialized individuals are, in almost all cases, worse than imported individuals from the elitist archive.

When only the elitist archive is imported, and the initialization of the population is done randomly, after a few generations, all the individuals in the population will be replaced by indi-

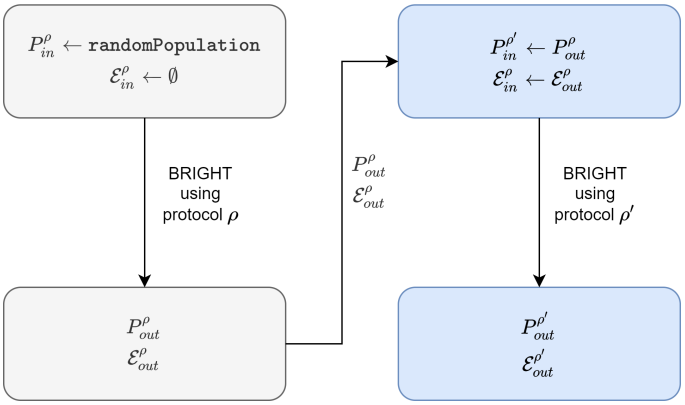


Figure 4.2: Structure of execution for importing the population and/or elitist archive in order to improve convergence speed of the second run. The notation is defined in Section 3.1.1

viduals from the elitist archive. Therefore importing the elitist archive will only be tested while also importing the population. The structure is visualized in Figure 4.2

5

Predicting the changes in dwell times

Besides importing the population and elitist archive, there might also be other ways to improve the convergence time after making a change in the protocol. In this chapter, the relation between the plans (individuals) of the output of BRIGHT between a run on the original clinical protocol and a run on the modified clinical protocol will be compared. When there is a clear relation, based on the modification of the clinical protocol, this relation could be exploited in order to create a model which could predict the changes in the dwell times between the results based on the original protocol and the results based on the modified protocol. This way the 'warm-start' could get a boost to be even more effective.

5.1. Relation

The goal of this research is to find a way to improve the convergence speed after changing the protocol (from ρ to p'). P^{ρ}_{out} and $\mathcal{E}^{\rho}_{out}$ are available to use, the goal is to reach the output plans of BRIGHT $P^{\rho'}_{out}$ and $\mathcal{E}^{\rho'}_{out}$. Importing P^{ρ}_{out} and $\mathcal{E}^{\rho}_{out}$ as initial population and initial elitist archive for the second run on ρ' might help, but there might be ways in order to push the result closer to $P^{\rho'}_{out}$ and $\mathcal{E}^{\rho'}_{out}$ than just importing P^{ρ}_{out} and $\mathcal{E}^{\rho}_{out}$.

When the change from ρ to ρ' is not too significant, it is likely that P^{ρ}_{out} and $P^{\rho'}_{out}$ are similar, even when the initialization of the run on ρ' is done using a random initialization ($P^{\rho'}_{in}$ = randomPopulation and $\mathcal{E}^{\rho'}_{in} = \emptyset$). Besides P^{ρ}_{out} and $P^{\rho'}_{out}$ being similar, the differences between those two sets of plans also might be predictable, based on the difference between p

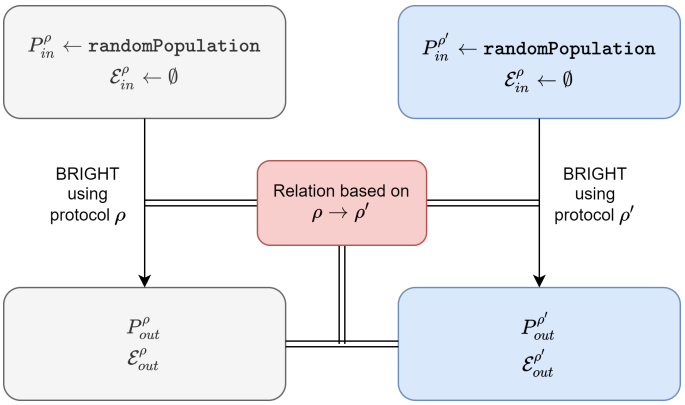


Figure 5.1: Relation between dwell times of the plans of the initial run compared to the run on the changed protocol

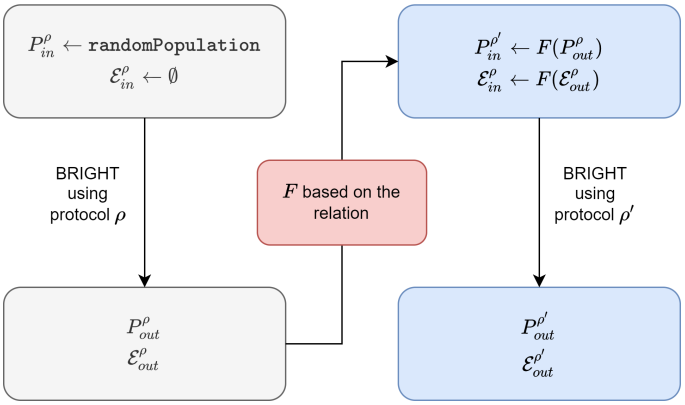


Figure 5.2: Usage of a model to predict the changes in dwell times

and p'. P^{ρ}_{out} and $P^{\rho'}_{out}$ will have different plans because both are optimized based on a different protocol. The differences between P^{ρ}_{out} and $P^{\rho'}_{out}$ will be related to the difference in the protocol (difference between ρ and ρ'). For example, when $D^{bladder}_{1cm^3}$ is made more difficult to achieve in ρ', then dwell positions in close proximity of the bladder are more likely to have lower dwell times for $P^{\rho'}_{out}$ compared to P^{ρ}_{out}.

However, the relation between P^{ρ}_{out} and $P^{\rho'}_{out}$ is based on the change which is made from ρ to ρ' might not be as straightforward. For example, when taking the previous example of decreasing the aspiration value value of $D^{bladder}_{1cm^3}$ (therefore making it more difficult to achieve), it is likely that the dwell positions will have lower dwell times around the bladder in $P^{\rho'}_{out}$. However, since it is a multi-objective optimization with a clear trade-off between coverage and sparing, it is also likely that in order to achieve the coverage indices, that other dwell positions will have higher dwell times in $P^{\rho'}_{out}$ in order to compensate for the lower radiation around the bladder.

The goal is to reach $P^{\rho'}_{out}$ using P^{ρ}_{out}. Therefore, if we find the relation between these two sets of plans, we can exploit this relation by modifying P^{ρ}_{out} to be more like $P^{\rho'}_{out}$. Then P^{ρ}_{out} can be modified by a function $F(x, y)$, where x is the set of plans and y is the modification of the protocol (ρ'). When the function F is used, the result, $F(P^{\rho}_{out}, \rho')$ will then be more similar to $P^{\rho'}_{out}$ than P^{ρ}_{out}. Given this modification, converging to high-quality treatment plans may take less time compared to random initialization.

In the next sections, three different ways of determining function F will be explained. In Chapter 6 the results of these implementations will be shown.

5.2. Distance to organ of protocol change

The first way in which the plans will be modified before importing (function F) will be to change the dwell times based on the distance to the organ in which the dwell times are changed.

The main idea of this method is to, after a change in one of the indices, modify the dwell times in such a way that the changed index will perform better compared to no modification of the dwell times. In the case of a sparing index, the LSI will decrease (in case the change is not within the slack) when evaluating based on the modified protocol ($P^{\rho'}_{in} \leftarrow P^{\rho}_{out}$), when applying the method to the population which will be imported ($P^{\rho'}_{in} \leftarrow F(P^{\rho}_{out}, \rho')$), then the LSI is still expected to decrease, but since some of the dwell times will be lowered, the decrease in LSI is at most equal to the original decrease, but likely to be less. The same holds for increasing the difficulty in coverage indices. However, instead of lowering the dwell times, the dwell times will be increased.

For example, when the aspiration value value for the sparing index $D^{bladder}_{1cm^3}$ is decreased (meaning the bladder should receive less radiation, therefore making it harder to achieve), then

with this method, the dwell positions which are close to the bladder will be decreased. The reason for doing this is that P^{ρ}_{out}, based on the unmodified protocol, is optimized towards the original aspiration value value for $D^{bladder}_{1cm^3}$. When directly importing the population and evaluation on the changed protocol ($P^{\rho'}_{in} \leftarrow P^{\rho}_{out}$) where the protocol change is increasing the difficulty to achieve $D^{bladder}_{1cm^3}$, then $P^{\rho'}_{in}$ will have lower LSI values compared to P^{ρ}_{out} because of the change in $D^{bladder}_{1cm^3}$. Using this method, the decrease in LSI values is expected to be less substantial since the dwell times around the bladder are decreased. This results in $P^{\rho'}_{in} \leftarrow F(P^{\rho}_{out}, \rho')$, where it is expected that the modified version of the population ($F(P^{\rho}_{out}, \rho')$) is closer to the final result of the run on the modified protocol ($P^{\rho'}_{out}$) than just using the P^{ρ}_{out}.

There are two expected disadvantages of using this method in order to improve the convergence speed of a second run on a modified version of the clinical protocol.

The first disadvantage is that for changes to coverage indices, it guarantees that the LCI value after applying the function will either be equal or better than the LCI values of the unmodified output of the initial run ($F(P^{\rho}_{out}, \rho')$ has equal or better LCI values than P^{ρ}_{out}, also when evaluating based on ρ'), but the LSI values of the plans will either be worse or as good since the dwell times are decreased. The decrease of the LSI values could be rather substantial because the function does only take into account which dwell positions are close to the organ of which the aspiration value value is changed and does not take into account the distances to the organs which have coverage indices. This could be accounted for but makes modifying a set of plans without evaluations substantially harder. The aim of the modification is also not to optimize the plans in such a way that the plans can be used directly but more to give the second run on the modified protocol a head start in order to decrease the convergence time.

The second disadvantage is that when using a function to increase the dwell times, if a coverage index is increased in difficulty, it is expected that the difference between P^{ρ}_{out} and $P^{\rho'}_{out}$, is that $P^{\rho'}_{out}$ will have relatively higher dwell times for dwell positions which are in close proximity of the organ of which the coverage aspiration value was modified. However, since the plans are a trade-off between sparing and coverage, it is likely that the overall dose distribution might also shift in order to compensate for the increase in radiation near the subjected organ in order to spare the organs at risk. But the method defined above will only increase dwell times if a sparing index is increased in difficulty. This means that even though $F(P^{\rho}_{out})$ will be more similar to $P^{\rho'}_{out}$ than P^{ρ}_{out}, $F(P^{\rho}_{out}, \rho')$, $F(P^{p}_{out}, \rho')$ will not have decreased dwell times in order to make up for the lack of sparing.

This method has various ways of being implemented and has, based on the way of implementing various parameters, the implementation of this method will be explained in Chapter 6.

5.3. Modelling changes

Another way to approach defining a function F which modifies a set of plans in order to improve convergence time is to find a model which maps plans from P^{ρ}_{out} to $P^{\rho'}_{out}$.

This function needs to provide, given a change of an aspiration value of one of the DVIs, the changes which are likely to occur between P^{ρ}_{out} to $P^{\rho'}_{out}$. This mapping can then be used in function F to improve the plans in P^{ρ}_{out} in order to become more similar to $P^{\rho'}_{out}$. If the input of a second run is more similar to the expected result, then it is expected that the time required to converge is less on the second run.

The main disadvantage of this method is that the relation between P^{ρ}_{out} and $P^{\rho'}_{out}$ is unknown initially and should be learnt from existing runs, which have P^{ρ}_{out} and $P^{\rho'}_{out}$ as result. This means that for every protocol change ($\rho \rightarrow \rho'$) the sets of plans P^{ρ}_{out} and $P^{\rho'}_{out}$ should be known, which requires at least one execution. Since there are many changes in protocol possible, it is infeasible to do these runs for individual cases. However, if per DVI increase/decrease of the aspiration value a pattern can be found, then function F can use this pattern in case one of the DVI changes in order to get $F(P^{\rho}_{out}, \rho')$ more similar to $P^{\rho'}_{out}$. But since anatomy and cathether placement might

vary a lot from patient to patient, it is unlikely that a single mapping can be made for each DVI which then can be used on a different patient to the one from which the mapping was learned.

Another important thing to consider is that not all plans on a front will be affected in the same way when looking at the differences between P_{out}^{ρ} and $P_{out}^{\rho'}$. Some plans in the front will not be affected since the change in DVI does not influence the LCI or LSI value because the change is within the slack of the plan, whereas for some other plans, the modification has a more substantial effect outside of the slack of the DVI for the plan and thus changing the LCI or LSI value. This will also be considered for the implementation of function F. In the upcoming sections, the way in which these plans will be matched (from P_{out}^{ρ} to $P_{out}^{\rho'}$) will be explained.

As stated, the relation can be used to increase the similarities between $F(P_{out}^{\rho}, \rho')$ and $P_{out}^{\rho'}$. This relation should be extracted from fully converged runs on both the original and modified protocol (P_{out}^{ρ} and $P_{out}^{\rho'}$).

The way in which this relation will be explored is by matching the plans based on a similarity metric and then tracking the differences based on the matched plans. This will be further elaborated on in Section 5.4.

5.4. Plan matching

The main idea of plan matching is that whenever a set of plans is used as input for the second run on a changed protocol ($P_{in}^{\rho'} \leftarrow P_{out}^{\rho}$), that plans of P_{out}^{ρ} will become more similar to $P_{out}^{\rho'}$ during each generation of BRIGHT. However, since plans on the front that are similar in LCI/LSI values might not have similar dwell times, it is important to compare plans of the two fronts which are similar.

It is unfeasible to simply tag or track the plans which are imported ($P_{in}^{\rho'} \leftarrow P_{out}^{\rho}$) and find the relation between P_{out}^{ρ} and $P_{out}^{\rho'}$ by comparing how the plans did change over time. The reason for this is that the changes in the plans are rather significant and are too different from being considered the same plan.

Whenever a relation between plans of P_{out}^{ρ} and plans of $P_{out}^{\rho'}$ is found, this relation can then be used in order to modify P_{out}^{ρ} using the function F to improve the convergence speed of the second run. All of this is given a specific change in the clinical protocol since the relation between P_{out}^{ρ} and $P_{out}^{\rho'}$ is dependent on the change in the clinical protocol ($\rho \rightarrow \rho'$).

Again, since anatomies of patient and catheter placements can differ from case to case, it is important to mention that if a relation is found for a specific change for one patient that it is likely to only work for that specific patient. However, if the relations show similarities among different patient, there might be a possibility to define a more general relation that can be used for more than one patient.

One possible way to find plans that match is to find a one-to-one mapping based on a similarity score. However, the similarity measure can not solely be based on the combination of LCI/LSI values. The reason for this is that based on a change in one of the DVIs, either the LCI or LSI values will change, meaning that if plans would be matched on similar LCI and LSI values that for instance, if a sparing index is changed to be harder to achieve, then most LSI values will be different between P_{out}^{ρ} and $P_{out}^{\rho'}$. Since when modifying a sparing index, none of the coverage indices will be affected, the LCI of the plans will not change if being evaluated on ρ' (assuming the same DCPs are used).

It is also likely that when trying to find the relation between P_{out}^{ρ} and $P_{out}^{\rho'}$ the LCI value could be useful to find a matching of plans when ρ' is a modification of a sparing index. Likewise, when the aspiration value of a coverage index is modified, the LSI value could be used when trying to find matches between the plans. The reason for this is that it is expected that the worst-performing index of the coverage or sparing indices will change less substantially when a sparing or coverage index is changed, respectively.

The main disadvantage of matching based on LCI or LSI values is that it does not consider the characteristics of a plan. When considering the output of BRIGHT, which is a set of plans, it

can not be assumed that plans with similar LCI and LSI values are also similar. For this reason, it might also be interesting to look at the dwell times and try to find a relation between P_{out}^{ρ} and $P_{out}^{\rho'}$ using the dwell times.

When finding a matching of plans based on dwell times, the core idea is to find plans which are similar between P_{out}^{ρ} and $P_{out}^{\rho'}$, based on the characteristics of a plan. Two plans can have similar LCI and LSI values but different ways of achieving this. For instance, the LCI and LSI values can be determined by different, worse-performing indices. It is, however, hard to compare plans based on their characteristics, especially since the plans might have a lot of differences based on the amount the protocol has been changed.

The practical implementation of these different plan matching methods can be found in Chapter 6.

6

Experiments

6.1. Variables/settings

All experiments are executed on the same system (AMD Ryzen 5 5600X and a GeForce RTX 3070). The initial protocol for every run is the AUMC2020 protocol, which is described in Section 1.1.2. The population size of each run is 96 individuals (plans). For optimization, 100.000 dose calculation points, spread evenly across all organs, are randomly sampled in each of the organs. For the final evaluation, 500.000 random dose calculation points are used. None of the experiments were subject to a time limit. The stopping criterion was the number of generations (in general, 200 generations, which is considered to be converged based on previous experience). Based on an average of 100 runs, the average time for initialization (mainly interpolating the organs) was 50.51 seconds. The average time taken per generation (based on 100.000 dose calculation points) was 0.8772 seconds based on 100 runs of 200 generations.

Other variables for these experiments are; $\tau = 35\%$, the linkage tree is static, built based on the euclidean distance between the dwell positions beforehand, and $\lambda = 10$ (the exponential weight parameter for LCI and LSI), which are the same as previously used by Bouter et al., 2019.

For all the experiments, a random seed is set and saved for reproducibility for both the algorithm as also for the sampling of the dose points within the organs. When results from a run are used in another run (for instance, importing the population), the seed of the initial run for the algorithm and dose points is used.

For this research, the data of 23 patients were available who were previously treated at the Amsterdam UMC. These are numbered from MB01 to MB23. The final results, which are displayed in this document, are based on patients MB20, MB21, and MB22. MB20 and MB21 are similar to most other patients with regards to the LCI/LSI fronts after convergence and are picked because they are representative for a bigger group of patients. MB22 is chosen since it is an outlier and BRIGHT performs in a different manner when comparing with, for instance, the performance on patients MB20 and MB21.

6.2. Metrics

6.2.1. Generations/time until (clinical) convergence

The main goal of this research is to lower the time required for convergence when a change is made in the protocol.

Since this research is inspired by a use case from clinical practice (Section 2), it is important to measure the improvement of the plans with regard to the clinical relevance (as opposed to just looking at the statistical relevance of an improvement). However, the clinical relevance of an improvement in the front of plans is hard to quantify since the clinical quality of a plan (or a set of plans) can only be determined by an expert. But even when plans are observed by experts, different experts will have different opinions on whether a plan is good or not.

However, plans with better LCI and/or LSI tend to be better plans. Especially if a plan has positive LCI and LSI values, the plan satisfies the clinical protocol and therefore is generally considered to be better than a plan which does not satisfy those requirements. This is further explained in Section 6.2.2.

It is important to mention, with regards to the convergence metrics, that within the problem context, there is also an uncertainty based on the chosen number of dose calculation points (DCPs). This is because of the sampling of the dose calculation points combined with the delineation uncertainty. Because of these uncertainty factors, the clinical convergence is different to the statistical convergence. As stated in van der Meer et al., 2019, based on sensitivity analysis for a set of patients, the uncertainty on all indices except one is more than 1% when using 20.000 DCPs per organ. The set of patients and the settings used for BRIGHT might not be the exact same compared to this research, however, it can be reasonably assumed that these levels of uncertainty are also applicable to this study.

6.2.2. LCI/LSI fronts

The main way in which runs of BRIGHT will be compared is by the use of a LCI/LSI front. After execution, BRIGHT has optimized a set of plans, this set of plans will be visualized using plots with LCI on the x-axis and LSI on the y-axis. The main advantage of using these fronts as a metric is that it shows the two main objectives of BRIGHT, which are being optimized on.

When a plan satisfies an aim of an index within the protocol the score of an index is ≥ 0. Since LCI and LSI are determined by the worst performing coverage and sparing indices, it means that when LCI ≥ 0 all coverage indices are satisfied and when LSI ≥ 0 all sparing indices are satisfied. When a plan plan has LCI > 0 and LSI > 0, then it satisfies all indices, these plans are considered to be within the so-called golden corner. LCI/LSI fronts also allow a way to quickly spot whether there are plans within the golden corner.

6.2.3. L-score metric

Using the LCI/LSI fronts as defined in Section 6.2.2 gives a good insight into the convergence and LCI and LSI scores of a given moment in the optimization (in most cases, this is after termination). However, since it is assumed that BRIGHT is converged after 200 generations, meaning that the LCI/LSI front will look similar after 200 generations when running multiple times on different seeds. In order to see whether BRIGHT converges within fewer generations, the generations between 0 and 200 should be evaluated to see differences. The reason to use another metric in addition to the LCI/LSI front after termination is to get information about the progression of the LCI/LSI front over time. The main advantage of the L-value metric is that it gives an indication of how far the front is progressed towards/into the golden corner.

The L-value is a metric that gives insight into the progression of a front. The L-value of a plan can be found by taking the minimum of the LCI and LSI values. The L-plan is the plan with the highest L-value. The L-value of the L-plan (denoted as L) can be found in the following way.

$$F = \{(LCI_0, LSI_0), (LCI_1, LSI_1), \dots (LCI_k, LSI_k)\}$$

where (LCI_k, LSI_k) represents a plan (with index k) with its LCI and LSI value

$$L = \max(\{\min(LCI_0, LSI_0), \min(LCI_1, LSI_1), \dots \min(LCI_k, LSI_k)\})$$

It is important to mention that the L-value is based on the worst value of LCI/LSI of a single plan. However, BRIGHT tends to produce LCI/LSI fronts without gaps (especially for runs of more than 20 generations). If a front has a significantly better L-value than another front, it could be stated that the plans of the first front are in general better than the plans of the second front. If the L-value is positive, it means that there is at least one plan in the golden corner (which satisfies all the aspiration values). If the L-value is negative, it means that there are no plans that have reached the golden corner.

For visualization purposes, the L-value is converted to an L-score. The reason for not using the L-value for visualization is that the L-value will change rather fast for the first few generations

and less for the latter generations. This is the reason why a log-plot is preferred in order to also track convergence for the latter generations. However, this is not feasible to do when using the L-value since the L-value will likely start negative but will end up being positive if the front of plans reaches the golden corner. Because of this, the L-score is used. For each visualization, the values are re-mapped, the worst (in most cases negative) L-value will be mapped to 1, the best L-value will be mapped to 0.0001 (in order to ensure the visualization can be done using a log-plot), and the values in between are mapped in a linear way.

6.2.4. Significance difference in L-score

In this chapter, the L-score will be used to make several comparisons with regard to convergence. It is important to mention that because of the aforementioned re-mapping of the L-values that different L-score plots can not be compared amongst each other, since the maximum (1) and minimal value (0.0001) are mapped to the maximum and minimal values which occur in the data which is selected for the plot.

The L-score plots have a log scale on the y-axis. However, the L-values are mapped in a linear way. This allows for tracking the differences in convergence throughout the optimization since the difference in the L-value at the start is way more substantial than the difference after a few generations. A negative side-effect of this choice is the fact that even though two sets of plans can be very similar after convergence, it could still look like there is a substantial difference between the two sets of plans. Because of this reason, the main focus when looking at L-score plots will be on the progression of the L-score and not the difference between the two at the number of generations where it is assumed to be converged. For a comparison of converged fronts, the LCI and LSI values will be plotted.

6.3. Slack

As explained in Section 3.2, when evaluating a plan on the clinical protocol, the objective values (LCI and LSI) are mainly determined by only two indices. Because of this, the other indices have 'slack' in their aspiration value, meaning that these aspiration values can be changed without substantial influence on the LCI and LSI value of the plan (within the margins as described in Section 3.2.1). Slack is important because if an aspiration value is changed within its slack range after a run, then rerunning will not give substantially different plans compared to the run on the original protocol.

In order to find the slack of a plan, the formulas defined in Section 3.2 will be used. For this experiment, the slack value of a set of plans will be calculated. This gives insight into what can be changed within the protocol without substantial changes in the entire LCI/LSI front. To evaluate the slack for a set of plans, the slack for each index associated with each plan will be calculated. Then for each index, the worst-case (minimum) slack value is chosen, which will be the slack on that index for the set of plans.

The slack will be evaluated for a single plan (the L-plan), for all plans in the golden corner (all plans where LCI >0 and LSI >0), and for all plans.

The slack calculations are done based on averages of 30 runs, with 200 generations per run, based on the original protocol.

6.3.1. Results

There are various possible sets of plans for which the slack can be calculated (i.e., for one plan, for an entire front, for all plans in the golden corner), and there are several ways to visualize the amount of slack and which DVIs have slack. In this section, several ways of visualization will be shown.

The slack on the original protocol (ρ) for patient MB21 after running 200 generations of BRIGHT can be found in Figure 6.1. The output of BRIGHT consists of an approximation set of 714 plans. Out of the 11 DVIs, there are 4 DVIs for which none of these 714 plans is the worst-performing DVI for either the coverage or the sparing indices. This means that those 4 DVIs have some slack. However, since the minimum value of the slack for all 714 plans is taken, there is

not much slack. For example, for the $V_{200\%}^{prostate}$ the aspiration value was 15%, meaning that the 15% most irradiated volume of the prostate should at most receive 200% of the prescribed dose (30Gy). The slack for this DVI is -0.19, making the aspiration value with slack 14.81%, meaning that the 714 plans will not have substantial changes if the aspiration value is changed from 15% to 14.81% (it might have some influence, as described in Section 3.2.1, the second-worst sparing index determines at most 9% of the LSI, the change in the aspiration value is approximately $\frac{14.81\%}{15\%} \approx 1.27\%$, meaning that the LSI is influenced at most by $9.0\% * 1.27\% = 0.11\%$).

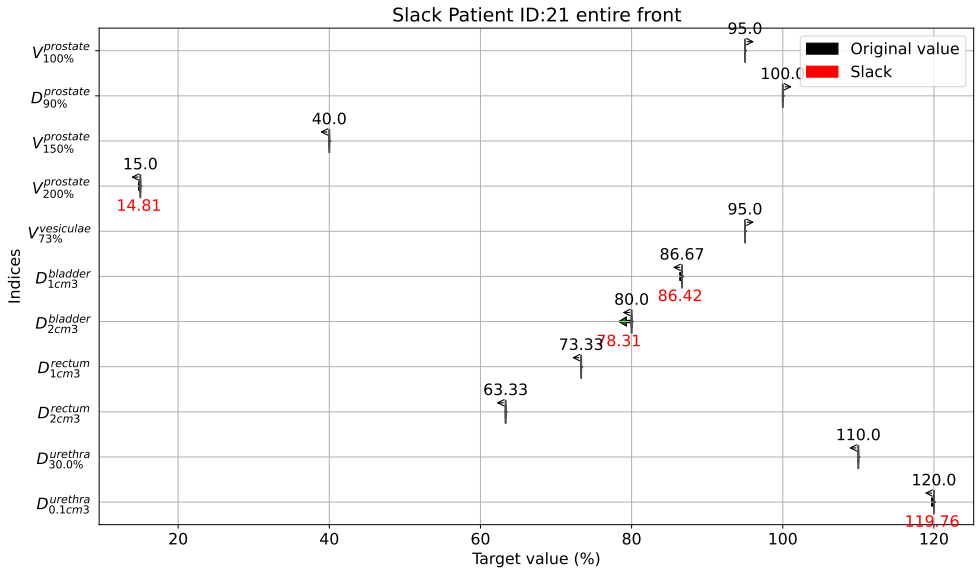


Figure 6.1: Slack for patient MB21. There is one row for each index, the vertical black line is the original aspiration value, and the little triangle indicates whether the index is a coverage index (pointing to the left) or a sparing index (pointing to the right). If there is an arrow from the black vertical line, then the index has slack. The new aspiration value with the slack is shown in red

In Figure 6.1 the absolute changes in slack can be seen. However, in order to analyze the amount of slack, a visualization showing the effect of the slack based on the percentual change manner might give better insights. For patient MB21 the diagram can be seen in Figure 6.2. Translating the amounts of slack to percentages also makes more sense when using these results in the upcoming section in which the aspiration values are changed based on a percentage of the aspiration value (Section 6.4), because these slack percentages can then be used without a conversion to compare with a percentual change in the aspiration value. This visualization also gives a better indication of the magnitude of the slack values based on their aspiration values.

Besides knowing the amount of slack, there is also valuable information to gain in which plans have zero slack based on the position on the LCI/LSI front (zero slack means most determining factor for either LCI or LSI). In Figure 6.3 the coverage indices analysis can be seen, and the number of plans for which an index is performing worst is indicated in the legend. It is important to mention that in some cases, two indices are the worst, instead of only one index. Because of this, the sum of the plans in the legend is greater than the number of plans in the front.

Figure 6.3 shows a clear pattern for patient MB21, namely, plans with relatively low LCI and relatively high LSI (plans which have lower dwell times in general) do tend do have their LCI determined by $V_{100\%}^{prostate}$. For most plans with a LCI greater than 0 and lower than 1 the LCI is mostly determined by $D_{90\%}^{prostate}$. Whenever LCI is equal to 1, the worst performing indices are $V_{73\%}^{vesiculae}$ and $V_{100\%}^{prostate}$ for every plan, this is because both aspiration values are achieved and their score being 100%, this means that no matter how much more radiation is applied, the

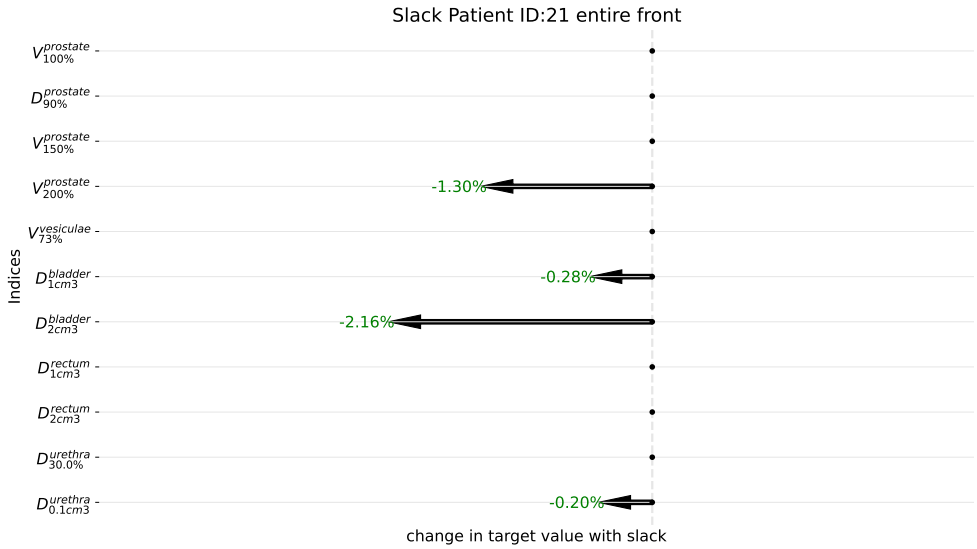


Figure 6.2: Slack for patient MB21 based on difference to the aspiration values. This diagram indicates the amount of slack related to the aspiration value. For example, when an aspiration value value is 50Gy, when the slack value is 49Gy, this diagram will indicate $\frac{49.0\text{Gy}}{50.0\text{Gy}} - 100\% = -2.00\%$

best score which can be reached is 100% of the volume of the organ. This is why $V^{vesiculae}_{73\%}$ and $V^{prostate}_{100\%}$ are both the worst performing indices for LCI ≈ 1.

So the main takeaway from Figure 6.3 is that for patient MB21, both $D^{prostate}_{90\%}$ and $V^{prostate}_{100\%}$ are the most influential parts of the LCI at different parts of the front. Within the golden corner (LCI and LSI >0), most plans have $D^{prostate}_{90\%}$ as worst performing index, however, especially for the low LCI values, $V^{prostate}_{100\%}$ is the worst performing coverage index within the golden corner.

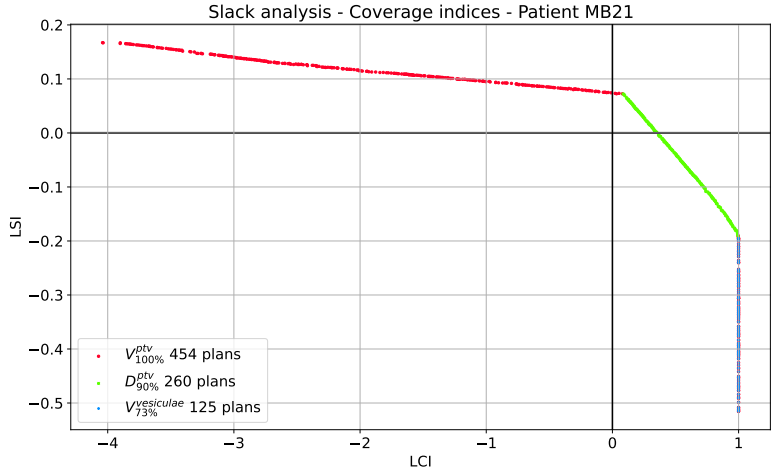


Figure 6.3: This front shows which coverage index is the worst performing index per plan (and therefore has a slack value of 0 and is the most important factor in determining LCI). The number of plans for which an index is performing worst is indicated in the legend.

The slack analysis with regards to the sparing indices can be seen in Figure 6.4. There are four indices which have zero plans which perform worst with regards to the sparing indices. These indices are $V_{200\%}^{prostate}$, $D_{1cm^3}^{bladder}$, $D_{2cm^3}^{bladder}$ and $D_{0.1cm^3}^{urethra}$. This means that based on the full front of 712 plans these four indices should have a slack value which is not equal to zero (since these indices are not the worst performing indices for any of the plans). This can be confirmed when looking at Figure 6.1.

Most plans, 557 of 712 have $D_{30\%}^{urethra}$ as worst performing sparing index. This likely means that for this specific patient (MB21) it is hard to satisfy the coverage indices while still sparing the urethra sufficiently. The plans which have $D_{30\%}^{urethra}$ as worst performing sparing index do span from LSI ≈ -0.3 to LCI ≈ 0.18. Only when the LSI is lower than -0.3 another index performs worse for the majority of the plans. The sparing index which performs worst for the majority of plans with LSI < -0.3 is the $V_{150\%}^{prostate}$, which indicates that the prostate is irradiated too much (the sub-volume of the prostate which is irriated with at least 22.5Gy is too big).

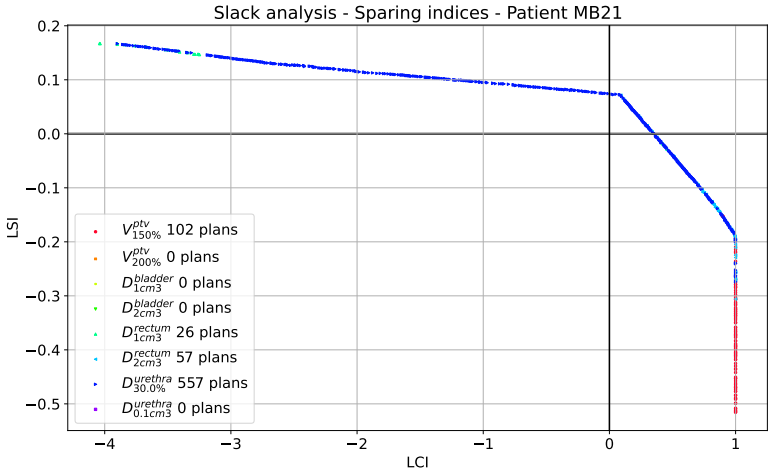


Figure 6.4: This front shows which sparing index is the worst performing index per plan (and therefore has a slack value of 0 and is the most important factor in determining LSI). The number of plans for which an index is performing worst is indicated in the legend.

The slack is, as expected, different from patient to patient. The main reason for this is because patients have different anatomies and the implants (catheters) might not be in the exact same positions. Most patients have fronts and also slack amounts which are similar to patient MB21 (results for these patients can be found in the Appendix (8)). However, there are some patients which have different fronts. For example, when the implants are positioned in such a way that makes it hard to reach the vesicles with radiation. An example of this happening is patient MB22. Because of this, this patient also has different amounts of slack. This can be seen when the slack of patient MB21 is compared to the slack of patient MB22. The slack diagram of patient MB22 can be seen in Figure 6.5. As shown in the diagram, there are three DVIs with slack, all are sparing indices. One of them, $D_{30\%}^{urethra}$ only has a minimal amount of slack (only 0.02% of the prescribed dose, which is 0.003Gy, while the other two indices, $D_{2cm^3}^{bladder}$ and $D_{1cm^3}^{rectum}$, have more slack, but also minimal compared to the slack for the other patient, MB21.

The fronts with worst performing DVIs can be seen in Figure 6.6. The results differ a lot compared to the fronts for MB21. The main important reason for this is that the implants of MB22 make it hard to reach the vesicles and therefore the aspiration value for $V_{73\%}^{vesiculae}$ is hard to satisfy. This can clearly be seen in Figure 6.6, since for a lot of plans (254 plans) is the worst performing coverage DVI. Plans for this patient also do not reach the golden corner (meaning no plan does satisfy the clinical protocol). Because of this there is no clear crossover between the plans with $V_{100\%}^{prostate}$ and $D_{90\%}^{prostate}$ which tends to happen for most patients between LCI values

around 0, in the golden corner.

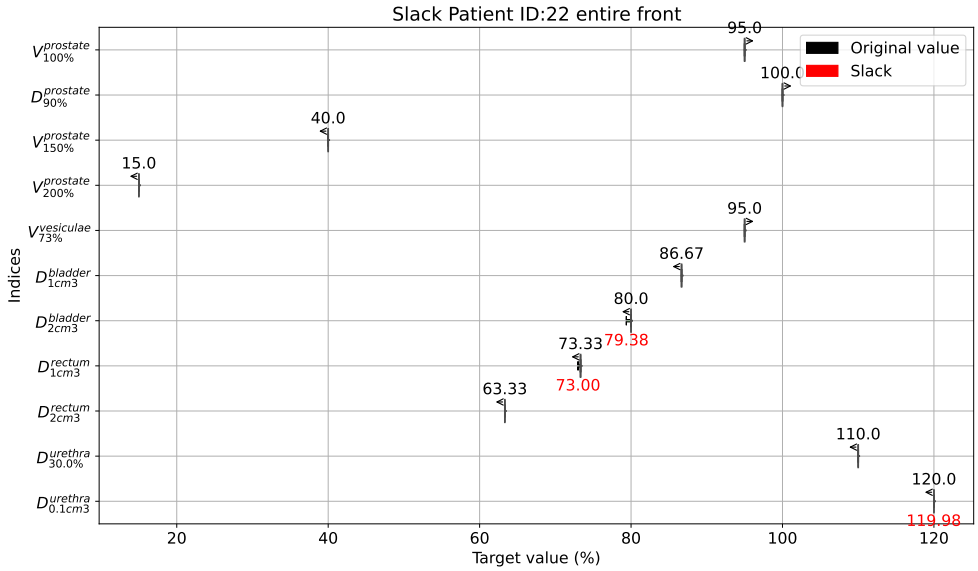


Figure 6.5: Slack for patient MB22.

The slack for other patients can be found in the Appendix (Chapter 8), but also for the other patients the slack is not substantial when considering the entire front. When translating the slack to a clinical use case, the the amount of slack is not a significant amount. In clinical practice the most interesting part of the front are the plans within the golden corner, as these plans do satisfy the protocol. Because of this it is also interesting to investigate the slack when only considering the plans within the golden corner.

Since the plans in the golden corner are a subset of all the plans in the entire front, the amount of slack will always be equal or greater when considering only the plans within the golden corner compared to the entire front.

The slack for patient MB21 when only considering the golden corner can be seen in Figure 6.7. As expected, the amount of slack on the DVIs either did not change or did increase (compared to Figure 6.7). Instead of only four DVIs which had slack, there are now seven DVIs with a slack value. Another observation which can be made is that $V_{150\%}^{prostate}$ and $V_{200\%}^{prostate}$ have a lot of slack when only considering the golden corner. The reason for this is that in order to satisfy the aspiration values for the other sparing indices (which is required in order to be in the golden corner) the radiation on around the prostate should be relatively low. Furthermore, this indicates that it is relatively easy to satisfy these constraints when all other sparing indices are satisfied.

There are four DVIs which have no slack, two of which are sparing indices and the other two are coverage indices. The slack analysis for both coverage and sparing indices can be found in Figure 6.8. As can be seen for the coverage indices the only two indices without slack are $V_{100\%}^{prostate}$ and the $D_{80\%}^{prostate}$. When looking back at Figure 6.3 the crossover point between $V_{100\%}^{prostate}$ and $D_{80\%}^{prostate}$ performing worst just lies within in the golden corner, which also can be seen when only considering the golden corner. With regards to sparing, the only two indices which have plans in the golden corner for which the index is the worst performing sparing index are the $D_{2cm^3}^{rectum}$ and the $D_{30\%}^{urethra}$, with mainly the $D_{30\%}^{urethra}$ determining the most of the LSI values of the plans, since it is the worst performing index for 79 plans. Some of the plans have both $D_{2cm^3}^{rectum}$ and $D_{30\%}^{urethra}$ as worst performing indices. When comparing with the slack analysis for the coverage indices, there is not really a cross over point between the worst performing indices, since worst performing indices for $D_{2cm^3}^{rectum}$ can be found at (within the golden corner) low LCI and high LCI values.

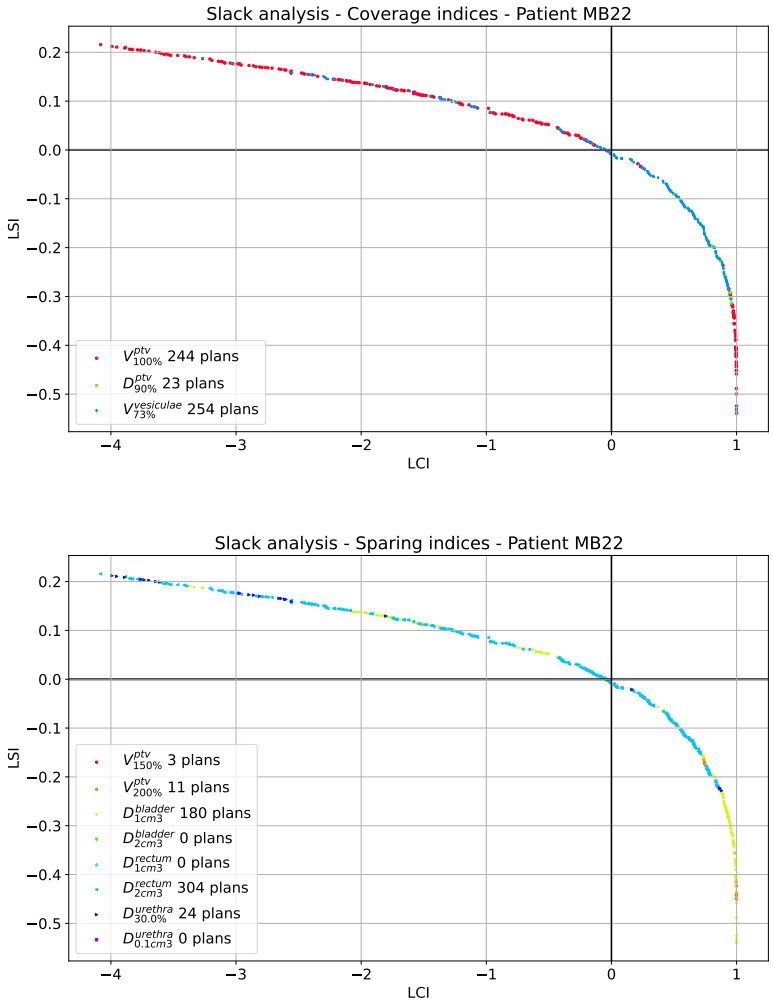


Figure 6.6: Worst performing DVIs for patient MB22.

However, in clinical use case, when an aspiration value is changed by an expert based on the initial results not being sufficient, the expert looks for different looking plans which do reflect the change in the protocol. Therefore it is most likely that the changes outside of the slack, which require further optimization, are more important. Within the next section the influence of changes outside of the slack will be evaluated.

6.4. Influence of changing the protocol

In order to improve the time it takes to converge after changing the protocol it is important to analyse what effect changing the protocol has on the results of BRIGHT. As stated before in Section 1.1.2, there are 3 coverage indices and 8 sparing indices. Each of these indices can have their aspiration value changed, so for each of these indices several changes of aspiration value will be tested. The goal is to find the effect of a modification within the protocol which is outside of the slack range for the modified DVI.

First a run on the original protocol will be executed, this run will be using a maximum of 200 generations (from which can be reasonably assumed that the front of plans has converged, based on previous experience). After this run a new run will be done on the modified protocol. The

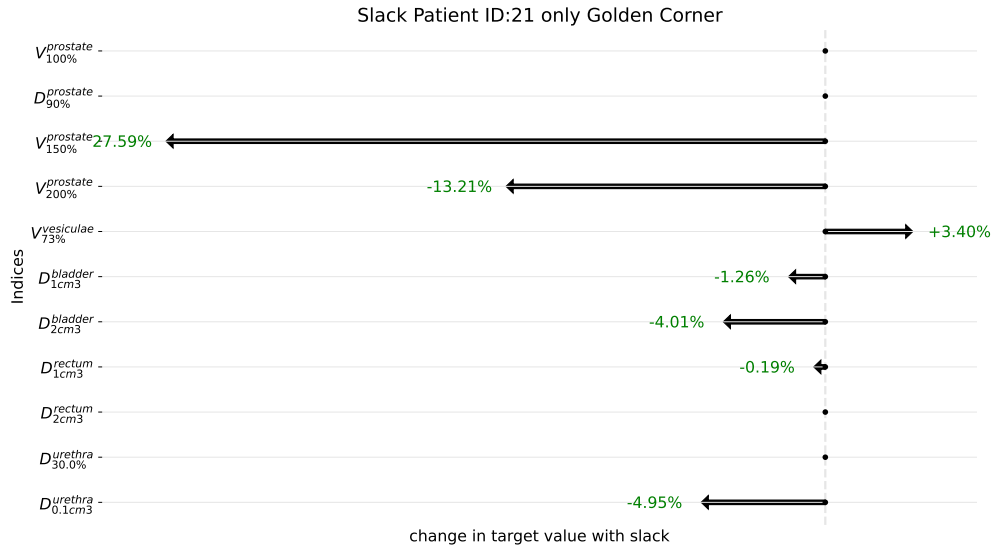


Figure 6.7: Slack for patient MB21 based on difference to the aspiration value. Only plans within the golden corner

fronts from these runs will be compared to see the influence of changing the specific aspiration value.

In order to deal with randomness and uncertainty of the runs, each experiment is executed and the LCI/LSI values of these runs are then plotted. Since performance of indices differs from patient to patient, the effect of the changes will be tested on different patients.

The main point of interest of this experiment is what the effect is on the convergence and the final front based on the LCI and LSI value. To better visualize the convergence L-value plots as described in section 6.2.3 will also be provided.

This specific experiment is conducted on 6 different patients. These are patients MB01, MB02, MB03, MB20, MB21 and MB22.

The influence of the protocol changes is visualized in plots with LCI on the x-axis and LSI on the y-axis, as explained in Section 6.2.2. For each index three changes which make it harder to achieve the index are applied. This means that the aspiration value of the specific index is increased or decreased with a factor. To make a coverage index harder the aspiration value will be increased (for example, increasing difficulty of $D_{90\%}^{prostate}$ >15Gy with 8% means that the aspiration value will be 15Gy $*1.08 = 16.2$Gy, making the index $D_{90\%}^{prostate}$ >16.2Gy). In order to increase the difficulty on a sparing index, the aspiration value will be decreased (for example increasing difficulty of $V_{150\%}^{prostate}$ <40% with 8% means that the new aspiration value will be 40% $*0.92 = 36.8\%$, making the index $V_{150\%}^{prostate}$ <36.8%). For each index three different changes in aspiration value will be evaluated.

6.4.1. Results

Not all results will be visualized in this section. The complete results can be found in the appendix.

First the results of the coverage indices will be evaluated. Then some of the sparing indices will be modified. All aspiration values will be increased in difficulty to achieve, meaning that, for the coverage indices, the aspiration value will be increased and, for sparing indices, the aspiration value will be decreased. For each of these indices the results for three different patients will be shown. These patients are MB20, MB21 and MB22, MB20 and MB21 seem to be similar to each other and the majority of the other patients with regards to the performance of BRIGHT,

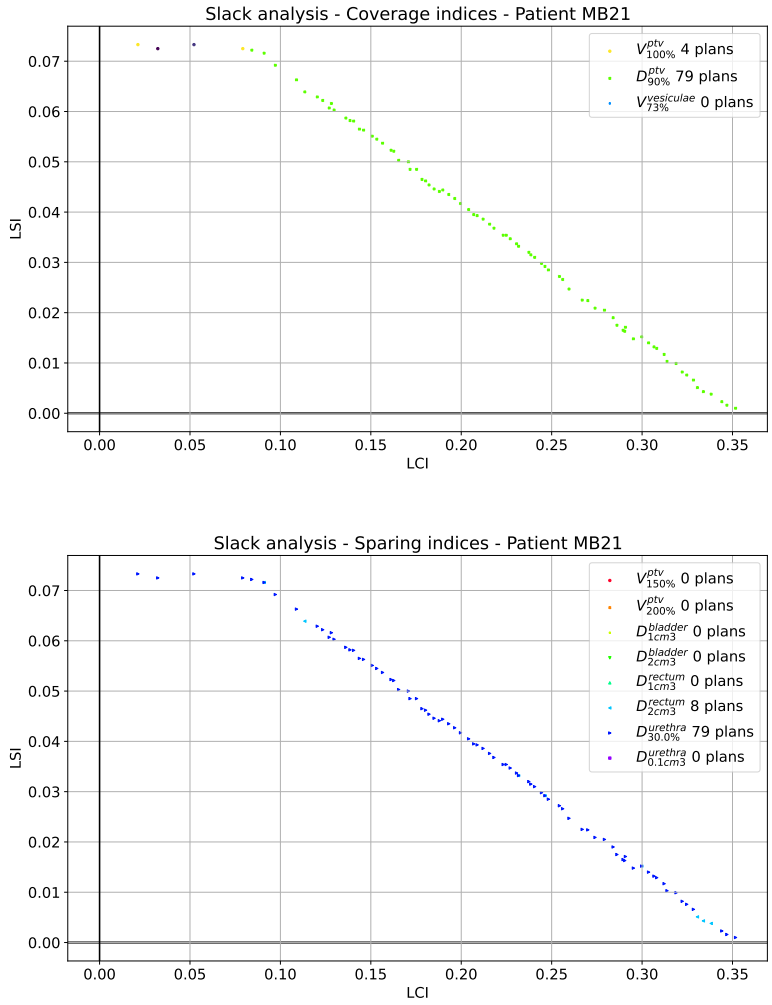


Figure 6.8: Slack analysis of plans in the golden corner for both coverage and sparing indices for patient MB21. The number of plans for which an index is performing worst is indicated in the legend.

whereas patient MB22 is different with regards that BRIGHT is not always able to find results in the golden corner after 200 generations.

$D^{prostate}_{90\%}$ >15Gy

The results of changing $D^{prostate}_{90\%}$ can be seen in Figure 6.9. As can be seen, the LCI value becomes worse after increasing the difficulty of the aspiration value for the $D^{prostate}_{90\%}$ index. This is to be expected since the change increases the difficulty of a coverage index. However, this decrease in LCI does not hold for all plans on the front, especially plans with an LCI of <-1.0 are mostly unaffected. The reason for only a part of the front changing is because for the unchanged parts of the front the $D^{prostate}_{90\%}$ is not the worst performing index for the LCI (and will not be the worst performing index for the LCI after changing).

For the 8% changes in most cases there is still some overlap with the plans of the non-modified AUMC protocol runs at the low sparing and high coverage part of the front (relatively low LSI, relatively high LCI), however this overlap is gone when the difficulty is increased with 16% or more. From this can be concluded that for these plans (low sparing, high coverage), 8% does not make it the worst performing coverage index, but 16% does make it the worst performing

coverage index. This is also to be expected when looking at the slack figures of Section 6.3, where the LCI/LSI fronts of MB20 and MB21 do bend, the plans have $D_{90\%}^{prostate}$ as worst performing coverage index. When comparing that to Figure 6.9, it can be seen that that specific part of the front is also the first which is influenced by a change of 8%.

When $D_{90\%}^{prostate}$ is changed with at least 16% the plans with a LSI value of 1 are also influenced. When comparing the coverage indices, the $D_{90\%}^{prostate}$ can go over 100%, this is since the requested unit is radiation on not a sub volume of an organ. However, the other two coverage indices, $V_{73\%}^{vesiculae}$ and $V^{prostate\,100\%}$ are limited to 100% of the organs volume, meaning that if $D_{90\%}^{prostate}$ is over 100% either $V_{73\%}^{vesiculae}$ or $V_{100\%}^{prostate}$ will be the worst performing coverage index. However, as can be seen, with a 16% change the $D_{90\%}^{prostate}$ becomes the worst performing coverage index for the plans which used to have an LCI of 1.

The results for patients MB20 and MB21 are similar, while the front for patient MB22 looks different. For patient MB22 BRIGHT barely reaches the golden corner (where LCI >0 and LSI >0). From this can be assumed that it is harder to find plans which adhere to the AUMC protocol with BRIGHT. But with regards to the change in the protocol it has a similar effect on the front compared to patient MB20 and MB21.

$V_{100\%}^{prostate}$ >95.0%

For $V_{100\%}^{prostate}$ only two modifications will be considered, the reason for this is that since the aspiration value is a percentage of an organs volume, that the maximum value for this index is 100%. Therefore only changes of 2% and 4% will be considered (meaning the aspiration value will be changed to 96.9% and 98.9% respectively). The effect of those changes can be seen in Figure 6.10.

The changes have similar effects on all three patients. With a change of 2% it is mainly the plans with LCI <0 which are affected, mainly for patient MB20 and MB21. For patient MB22 the change of 2% seems to also affect some plans with LCI >0, but also all plans with LCI <0 are affected.

When looking at the slack figures in Section 6.3, the plans where $V_{100\%}^{prostate}$ was the worst performing coverage index are mostly the plans with LCI <0, which, when comparing with Figure 6.10, this does match, since the plans with LCI <0 are affected the most.

Even though the shapes of MB20/MB21 and MB22 do not look similar, increasing the difficulty to achieve $^{prostate}_{100\%}$ seems to have similar influence on the fronts.

$V_{73\%}^{vesiculae}$ >95.0%

As with the coverage index $V_{100\%}^{prostate}$, for $V_{73\%}^{vesiculae}$ also only two changes will be considered. The reason for this is the same as with the $V_{100\%}^{prostate}$, namely that the maximum requested volume can not go beyond 100%.

The results can be seen in Figure 6.11. For these modifications there is quite a big difference between patient MB22 and patients MB20 and MB21. For patients MB20 and MB21 the change does not seem to have any effect. This seems to not match the conclusions of the slack calculations of Section 6.3, when for instance comparing these results with Figure 6.3. However, as can be seen, those plans are all plans with a LCI value of 1, meaning that the maximum score for the worst performing coverage indices is reached. Since the slack is based on the distance to the worst performing and both $V_{prostate}^{100\%}$ and $V_{vesiculae}^{73\%}$ both have 100% of the requested sub volume covered by the requested dose, both are then worst performing. However, that does not mean that increasing the difficulty of $V_{vesiculae}^{73\%}$ with 2% or 4% will directly influence the LCI value. The reason for this is that if the requested volume is increased from 95% to 96.9% or 98.9% and the actual coverage of those plans has 100% of the volume covered, in that case nothing will change with regards to the LCI.

For all other plans on the front for patient MB20 and MB21 it also does not seem that changing $V_{73\%}^{vesiculae}$ has much effect on the plans. This means that by achieving the other coverage indices ($D_{90\%}^{prostate}$ and $V_{100\%}^{prostate}$) $V_{73\%}^{vesiculae}$ is already satisfied.

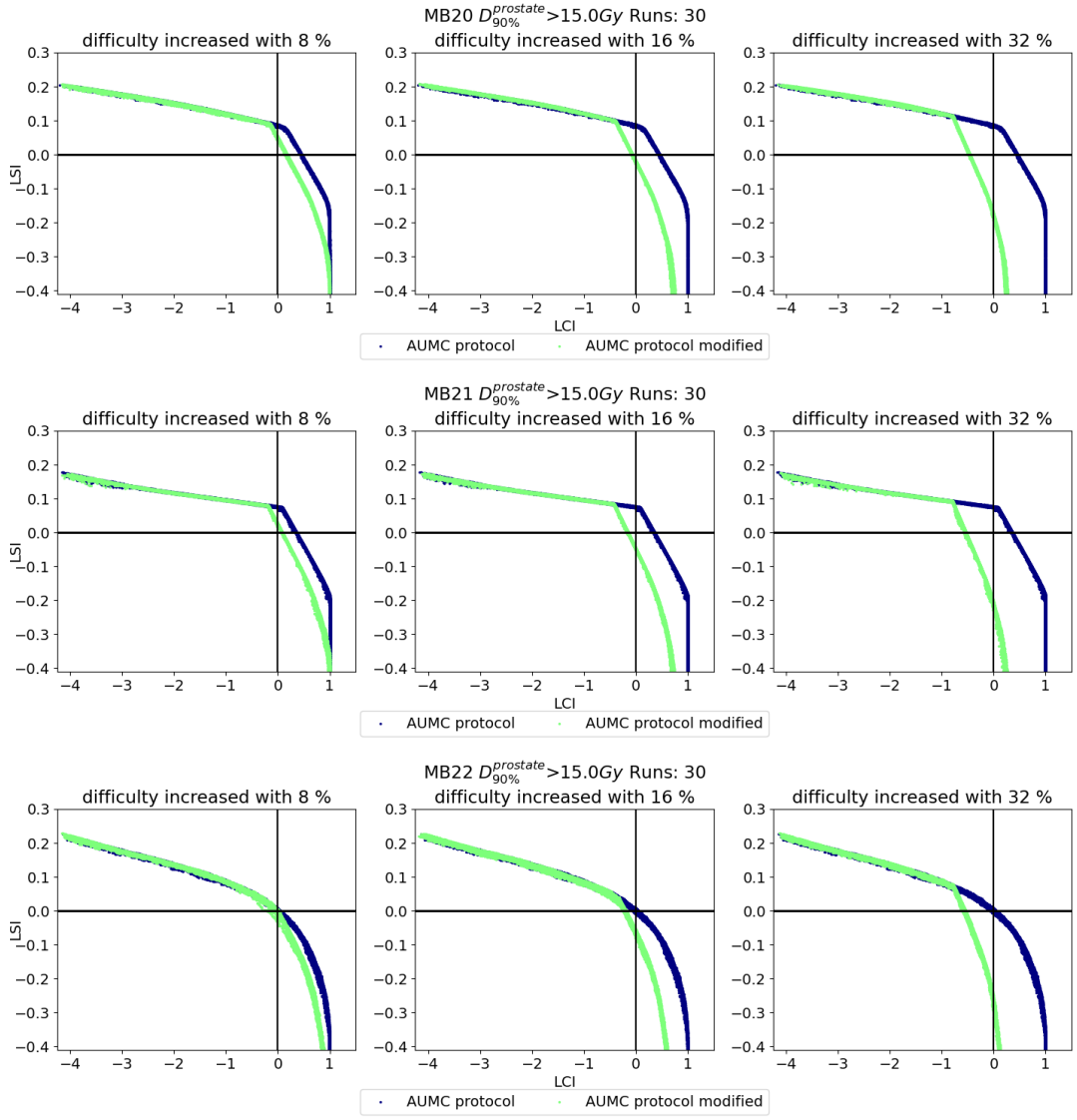


Figure 6.9: Influence of increasing the $D_{90\%}^{prostate}$ for patient MB20, MB21 and MB22 (note that the difficulty increase is only applied to the 'AUMC protocol modified')

The results for patient MB22 are very different. As can be seen in Figure 6.11, increasing the difficulty to achieve $V_{73\%}^{vesiculae}$ has an effect on the entire front. This also matches the slack results of Figure 6.6, where can be seen that a significant number of plans have $V_{73\%}^{vesiculae}$ as worst performing index. This shows that not only the fronts but also the performance of the different DVIs can vary a lot from patient to patient. For this patient, MB22, specifically, the locations of the dwell positions are sub optimal and make it hard to achieve the coverage required for the vesiculae.

$V_{150\%}^{prostate}$ <40%

The $V_{150\%}^{prostate}$ is a sparing index which aims to prevent applying too much radiation to the planning target volume. Since both $V_{150\%}^{prostate}$ and $V_{200\%}^{prostate}$ are similar sparing indices, but also have similar effects only changes in $V_{150\%}^{prostate}$ will be considered. The results can be found in Figure 6.12.

The slack of $V_{150\%}^{prostate}$ varies from patient to patient, but in general there was little to no slack.

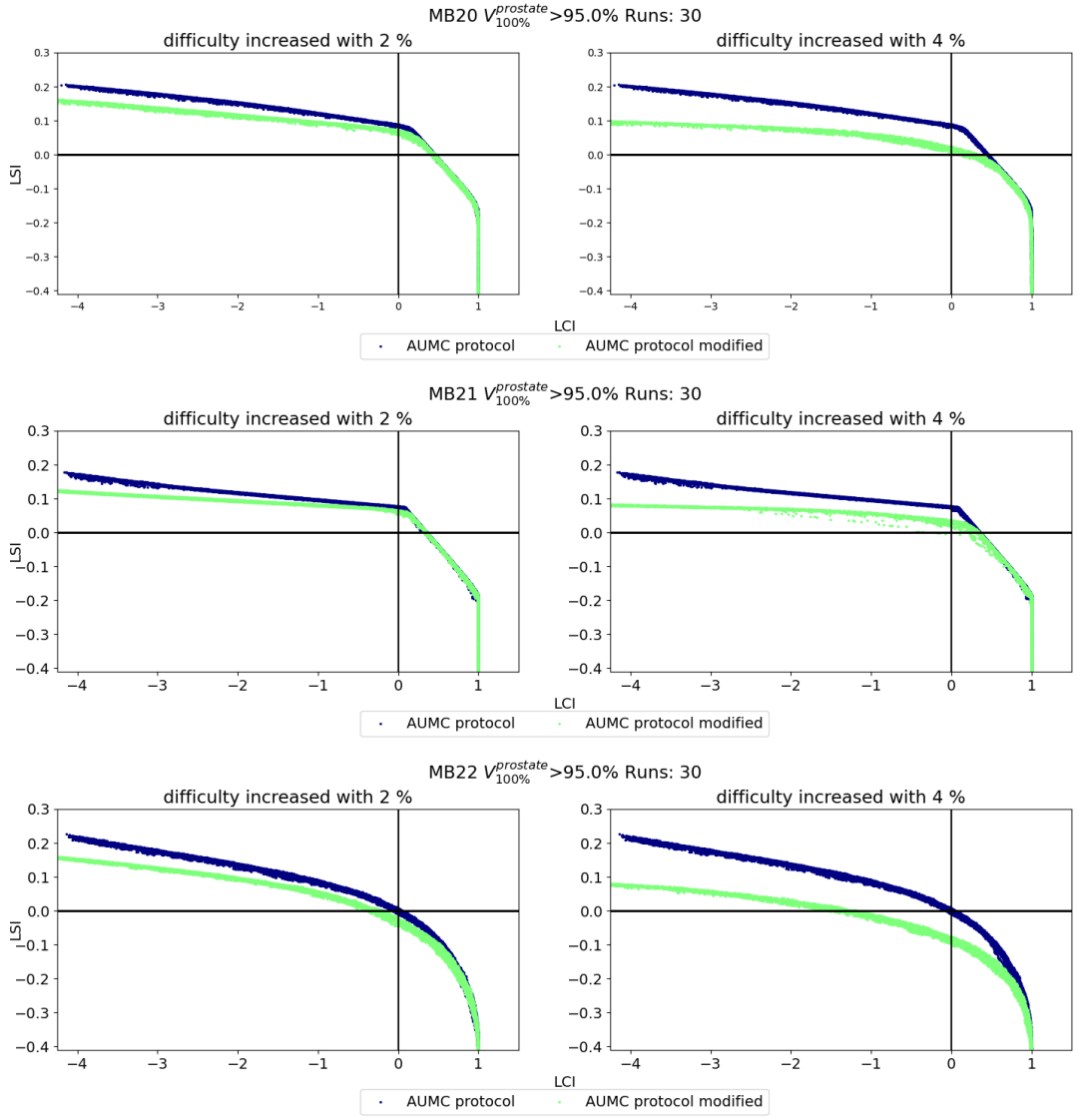


Figure 6.10: Influence of increasing the $V_{100\%}^{prostate}$ for patient MB20, MB21 and MB22 (note that the difficulty increase is only applied to the 'AUMC protocol modified')

In case plans had $V_{150\%}^{prostate}$ as their worst performing sparing indices, the plans mostly had an LCI of 1, meaning plans with relatively high radiation. The reason for this is that those plans have longer dwell times in general, meaning that there are dose points within the prostate with over 150% radiation.

The effect of changing $V_{150\%}^{prostate}$ with 8% (to 36.8%) has barely any effect on the resulting front, the change of 16% (to 33.6%) also shows similar front, with only slight changes. Finally, the increase of difficulty for $V_{150\%}^{prostate}$ with 32% (to 27.2%) does have more influence, it seems that especially for patient MB22 there are some plans on the front which have a lower LSI value.

$D_{1cm^3}^{bladder}$

The effects which a change of $D_{1cm^3}^{bladder}$ on patient MB20, patient MB21 and patient MB22 can be seen in Figure 6.13. The modification of 8% has little to no effect for patient MB20 and patient MB21, but seems to influence MB22 more. This is likely because of the fact that there is less margin because the aspiration value for $V_{73\%}^{vesiculae}$ is harder to achieve for patient MB22.

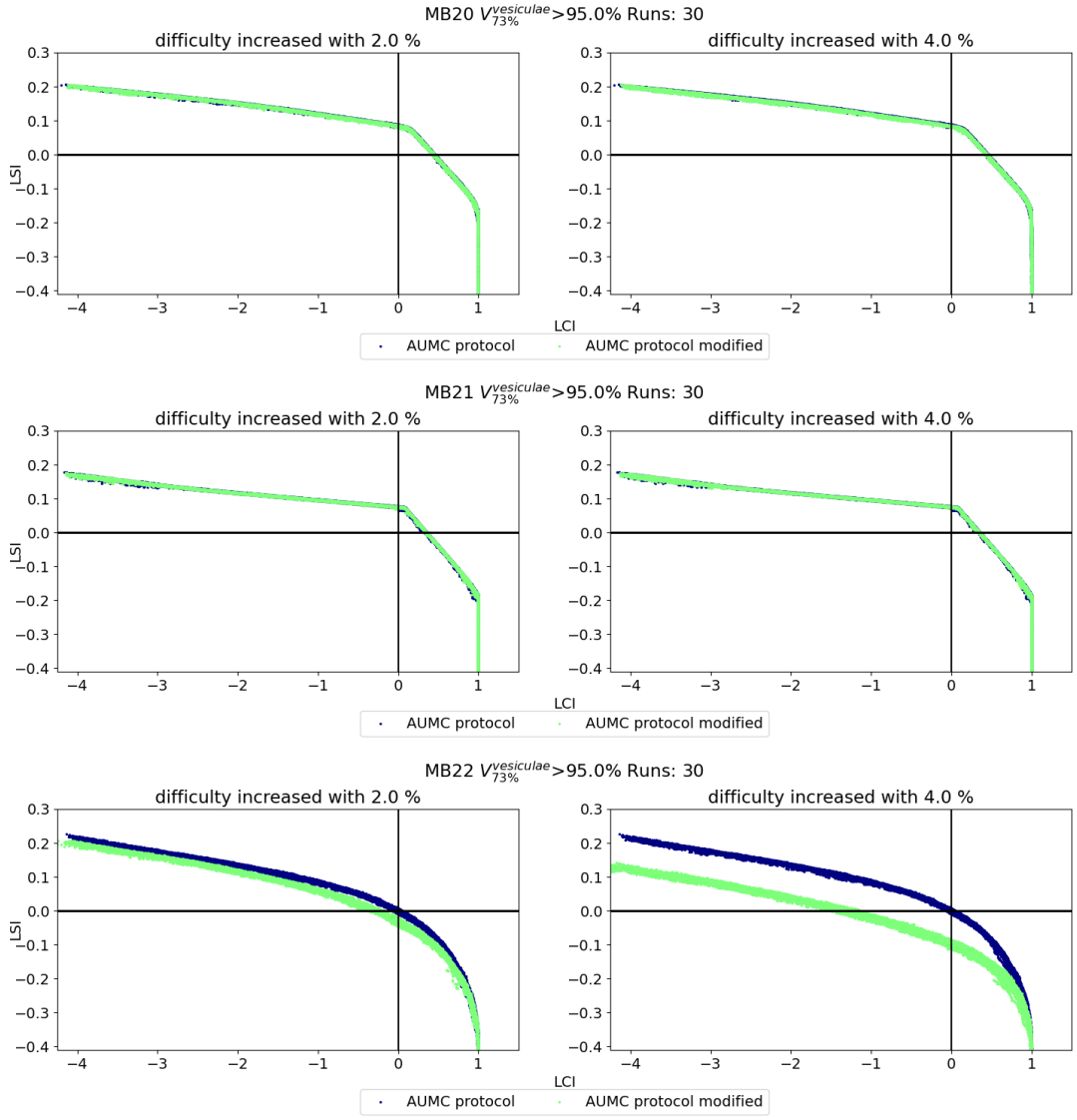


Figure 6.11: Influence of increasing the $V_{73\%}^{vesiculae}$ for patient MB20, MB21 and MB22 (note that the difficulty increase is only applied to the 'AUMC protocol modified')

When looking at 16% and 32% the entire front shifts because the LSI values are significantly influenced by the change in the protocol. This shows a similar effect for all three patients.

From these experiments can be concluded that changes to various DVIs have various effects. Even when changing the same DVI for two patients which show similar performance when being optimized with BRIGHT (patient MB20 and patient MB21), there are still differences in the effect of the change. From this can be concluded that the effect of a change can not be generalized. When looking at different percentual changes to the same DVI for a patient, there seems to be a pattern of how much the LCI/LSI values are affected.

6.5. Effects of importing individuals

In order to test the effect of importing the individuals experiments will be executed. As stated in Chapter 4, the population and the elitist archived will be imported ($P_{in}^{\rho'} \leftarrow P_{out}^{\rho}$ and $\mathcal{E}_{in}^{\rho'} \leftarrow \mathcal{E}_{out}^{\rho}$). In this section the experiments will be conducted, the main goal is to compare convergence speed

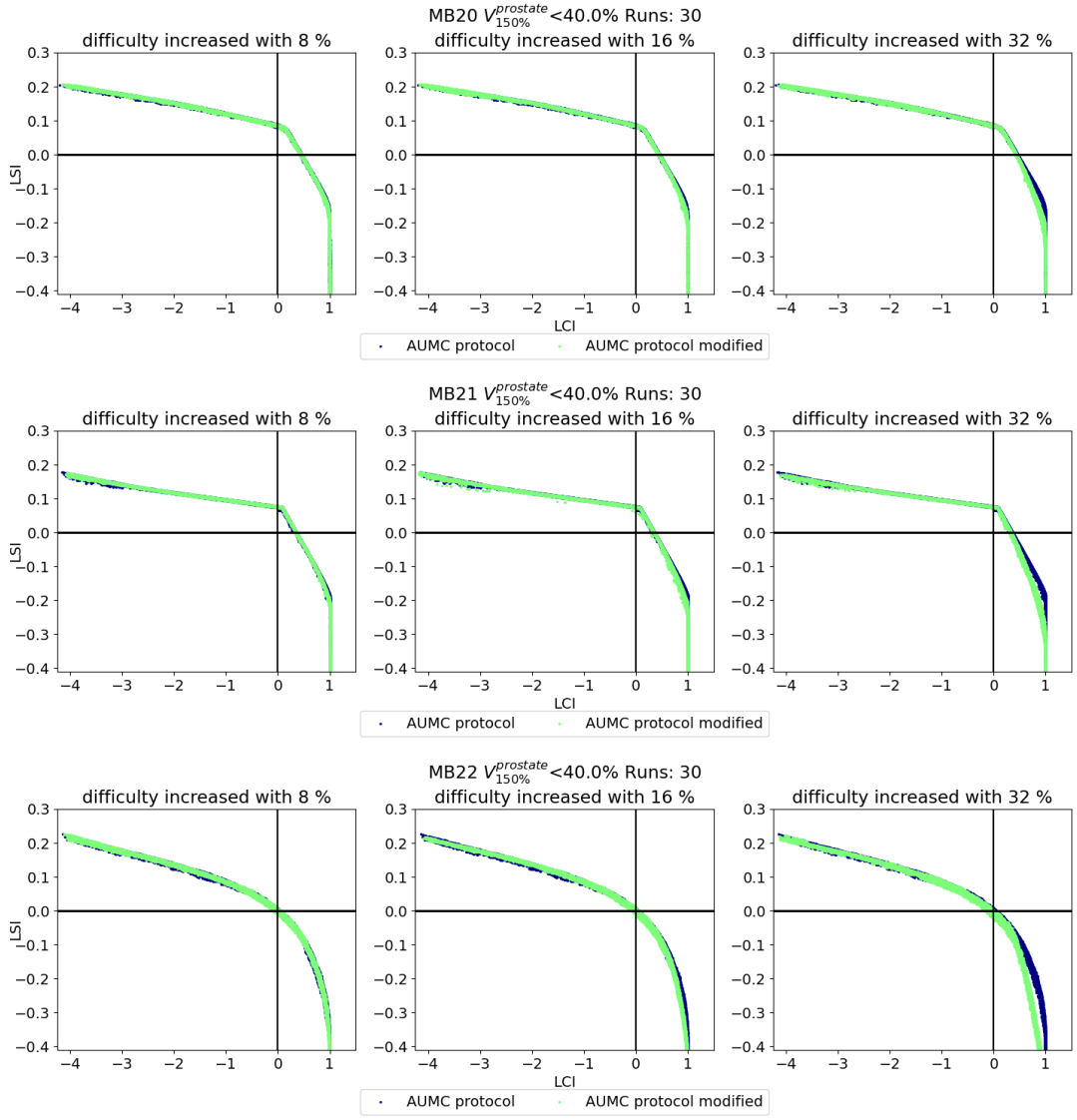


Figure 6.12: Influence of increasing the $V_{150\%}^{prostate}$ for patient MB20, MB21 and MB22 (note that the difficulty increase is only applied to the 'AUMC protocol modified')

between the default initialization compared to importing individuals.

It is important to make sure that when using imported population and/or elitist archive that the final results will be at least similar or better compared to random initialization. To check this 200 generations on the modified protocol will be compared between the random initialization and importing the population and the elitist archive.

First of all it is important to check whether the final result of starting with random initialization is similar to the final result of importing the population and the elitist archive. As explained in 4, importing individuals could possibly cause BRIGHT to be stuck in a local optima. Since 200 generations assumes that the front of plans is converged with a random population it is expected that the front of plans when importing the population and the elitist archive is also converged after 200 generations.

Secondly, if both settings (random initialization and importing population and elitist archive) give similar fronts after 200 generations, the focus will be on the speed of the convergence. As stated in the problem statement (in Chapter 2), the goal is to decrease the time required to

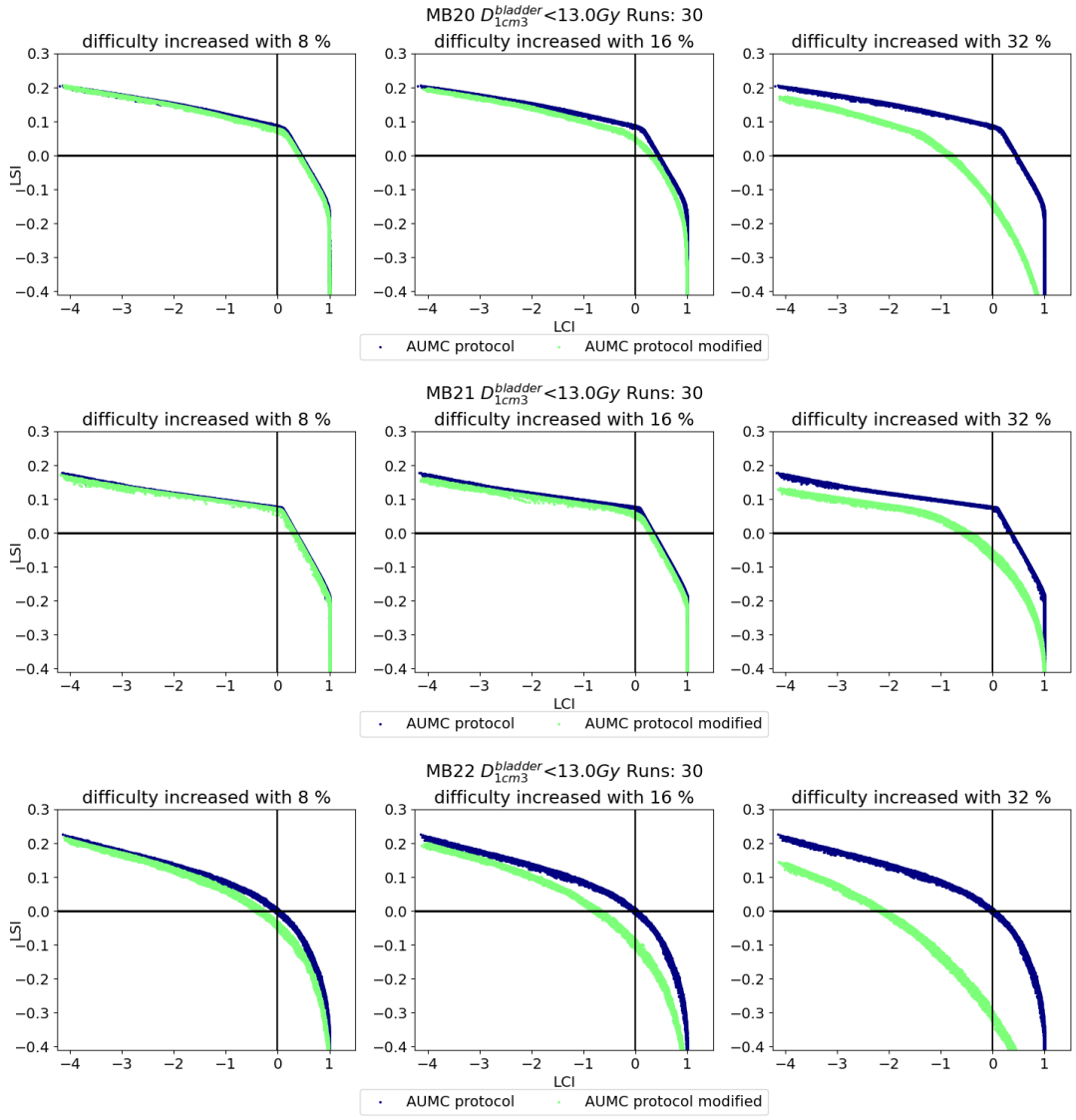


Figure 6.13: Influence of increasing the $D_{1cm^3}^{bladder}$ for patient MB20, MB21 and MB22 (note that the difficulty increase is only applied to the 'AUMC protocol modified')

converge to high-quality treatment plans after changing the protocol.

Figure 6.14 shows the results for patients MB20, MB21 and MB21 while changing $D_{90\%}^{prostate}$. As expected, the plans of the run on the original protocol perform better based on their LCI and LSI values, however, these should not be compared one to one as the objective space is different (because of the change in protocol). The results of the original protocol are shown for reference.

From Figure 6.14 can be concluded that for this case random initialization and importing the population and the elitist archive have similar results, for all three patients almost all points of 'AUMC protocol modified' do overlap with 'AUMC protocol modified - imported population and archive', there is so much overlap that in some cases there is barely any difference in the plots. With this can be concluded that importing the population and the elitist archive does not cause the population to be stuck in a local optimum after 200 generations for a modification of $D_{90\%}^{prostate}$ for patients MB20, MB21 and MB22.

Now that for an increase of difficulty of $D_{90\%}^{prostate}$ it shows that both random initialization and importing plans do converge to similar fronts, the convergence speed over time can now be

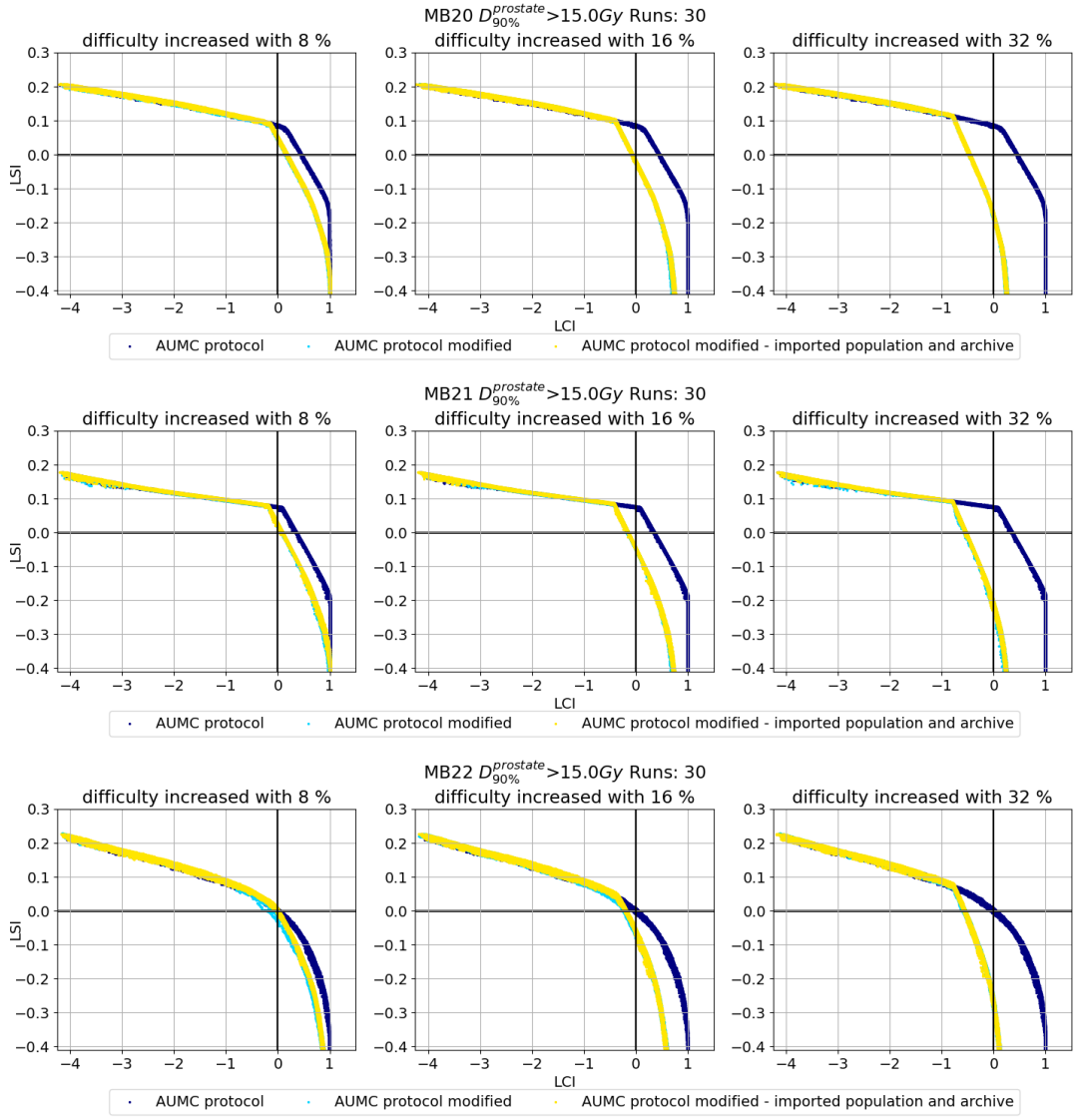


Figure 6.14: Modification of $D_{90\%}^{prostate}$ for MB20, MB21 and MB22. LCI/LSI fronts for the original protocol, the modified protocol with random initialization and the modified protocol with importing the population/elitist archive of an initial run

compared. Since the time required to converge could differ when running on different hardware, the time will be tracked in the amount of generations which BRIGHT has executed. In order to track the convergence over time the L-score will be used. As explained in 6.2.3, the L-score can be used in order to visualize convergence over time.

The L-score over generations for three different modifications of $D_{90\%}^{prostate}$ for patient MB20 can be seen in Figure 6.15. The y-axis ranges from 0.0001 to 1 as defined in Section 6.2.3 and the shaded areas are the 90th percentile of multiple runs, in this case 30 runs. In the Figure the L-scores of default initialization ($P_{in}^{\rho'} \leftarrow$ randomPopulation and $\mathcal{E}_{in}^{\rho'} = \emptyset$) are compared with the L-scores of importing the population and the elitist archive ($P_{in}^{\rho'} \leftarrow P_{out}^{\rho}$ and $\mathcal{E}_{in}^{\rho'} = \mathcal{E}_{out}^{\rho}$), where the modification in the protocol ($\rho \rightarrow \rho'$) is a modification of $D_{90\%}^{prostate}$. It seems like the uncertainty increases over the generations (shaded area), but this is not the case since the y-axis is a logarithmic axis. For all three plots it is clear that at the start, on zero generations, there is a

significant difference between the default initialization compared to importing the population and the elitist archive. This is to be expected, since with the default initialization all plans have their dwell times randomly initialized, whereas when the population and elitist archive are imported the plans will have all but one of their DVIs (the modified DVI) already optimized. As stated in Section 6.2.4, even though it seems like the default initialization does perform worse at 200 generations, because of the logarithmic axis this is not the case. This is also further shown by Figure 6.14, which is based on the same data.

Since there are a lot of similarities between patient MB20 and MB21, also with regards to the L-scores when applying a modification on $D_{90\%}^{prostate}$, patient MB21's L-scores are added to Appendix.

To see whether the effect on a different patient is similar the change of $D_{90\%}^{prostate}$ is also applied to patient MB22. For patient MB22 BRIGHT did not reach the golden corner within 200 generations and therefore a change might have different results. The L-score visualization of patient MB22 can be found in Figure 6.16. The results are mostly similar when compared to the results of patient MB20. The only main difference can be seen when looking at the modification of 32%, in which case it seems like the difference between the two different settings is less substantial. Meaning that in this specific case, for a change of 32% importing the population improves convergence speed more for patient MB20 than for patient MB22.

As the L-score plot is a metric which is based on the evaluation over generations of the L-plan it helps with tracking convergence speed, however, in order to confirm findings it is important to also look at the LCI/LSI fronts. In order to confirm the assumption based on L-score fronts that when changing $D_{90\%}^{prostate}$ with 32% MB20 benefits more from importing the population and the elitist archive than MB22 LCI/LSI plots at several points should be compared. When comparing Figure 6.15 and Figure 6.16, the difference mainly occurs from 75 generations onwards. Therefore this the fronts of 50, 75 and 100 generations will be compared. This is done in Figure 6.17. The approximation sets at 10, 30, 100 and 200 generations are visualized, and the run on the unmodified AUMC protocol is also shown as reference. When comparing to the L-score figures it can be seen that the main difference between 30 and 100 generations can also be spotted in the LCI/LSI fronts. At 30 generations for patient MB22 there is a significant difference between the default configuration and importing the population and elitist archive, which is gone at 100 generations, which also corresponds with the L-score plot. Besides that it is important to mention that, as described in Section 6.2.4, the L-scores could not be compared one to one between these two L-score plots. This is the reason why it seems like MB20 should show less similarity between both settings, but when looking at Figure 6.17 this is not the case.

When looking at Figure 6.15 it can also be concluded that for both patients, mainly for the first 30 generations importing the population and elitist archive positively influences the convergence speed. At 10 generations the front is already close to being fully converged when comparing the front with the front of 200 generations. For both patients there is also still a difference visible for 30 generations, although the difference for patient MB22 is more substantial, for patient MB20 most of the plans when importing the population and the archive still Pareto-dominate most of the plans when running using the default initialization.

6.6. Predicting the changes in dwell times

As stated in Chapter 5 another way which can be used in order to improve convergence speed is to find a way in order to improve the fitness of $P_{in}^{\rho'}$ and $\mathcal{E}_{in}^{\rho'}$ without optimization by BRIGHT yet. The methods explained in Chapter 5 will be evaluated in this section, these are modifying the dwell positions based on the distance to the organ of which the index changes, plan matching by matching based on LCI or LSI and plan matching based on similar plans.

6.6.1. Distance to organ of protocol change

The first way in which the changes in dwell times will be predicted is by changing the dwell times based on the approximate distance to the organ for which an aspiration value in the clinical

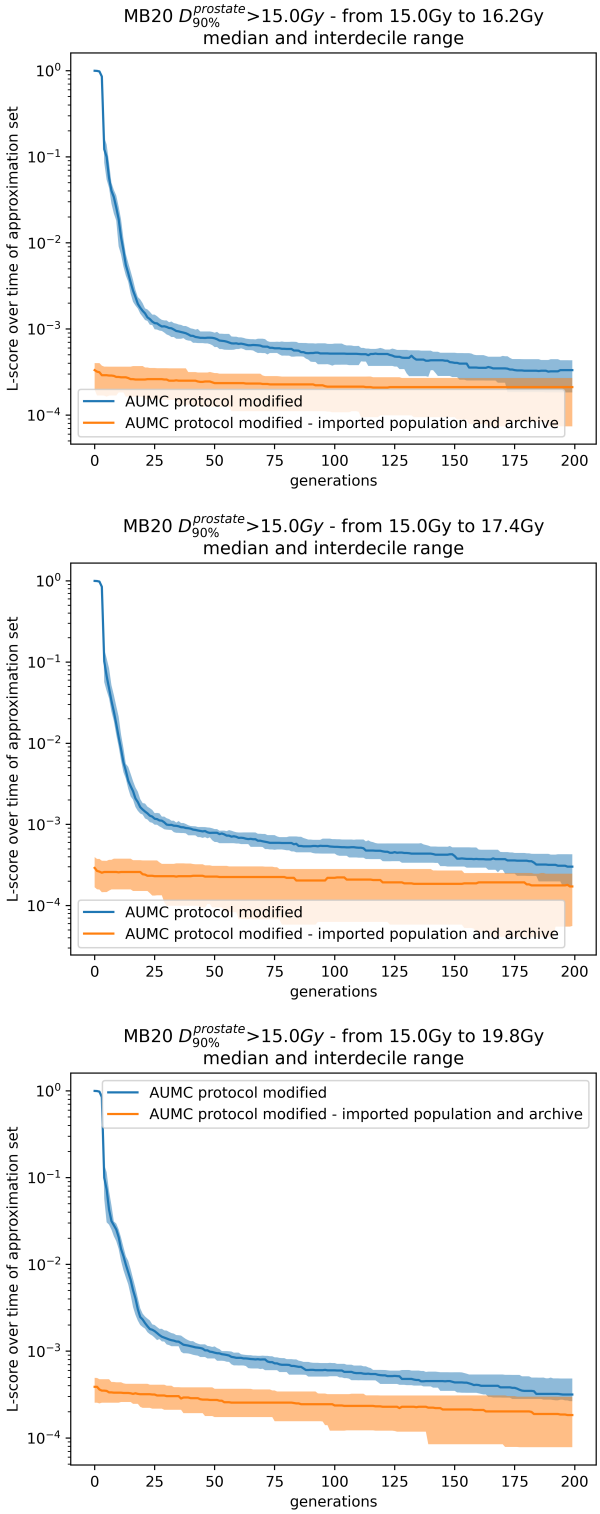


Figure 6.15: Median and interdecile range of the L-score for modifications of 8%, 16% and 32% on aspiration value of $D_{90\%}^{prostate}$ for patient MB20

protocol is changed. This is less of a modelled solution, but more based on an educated guess.

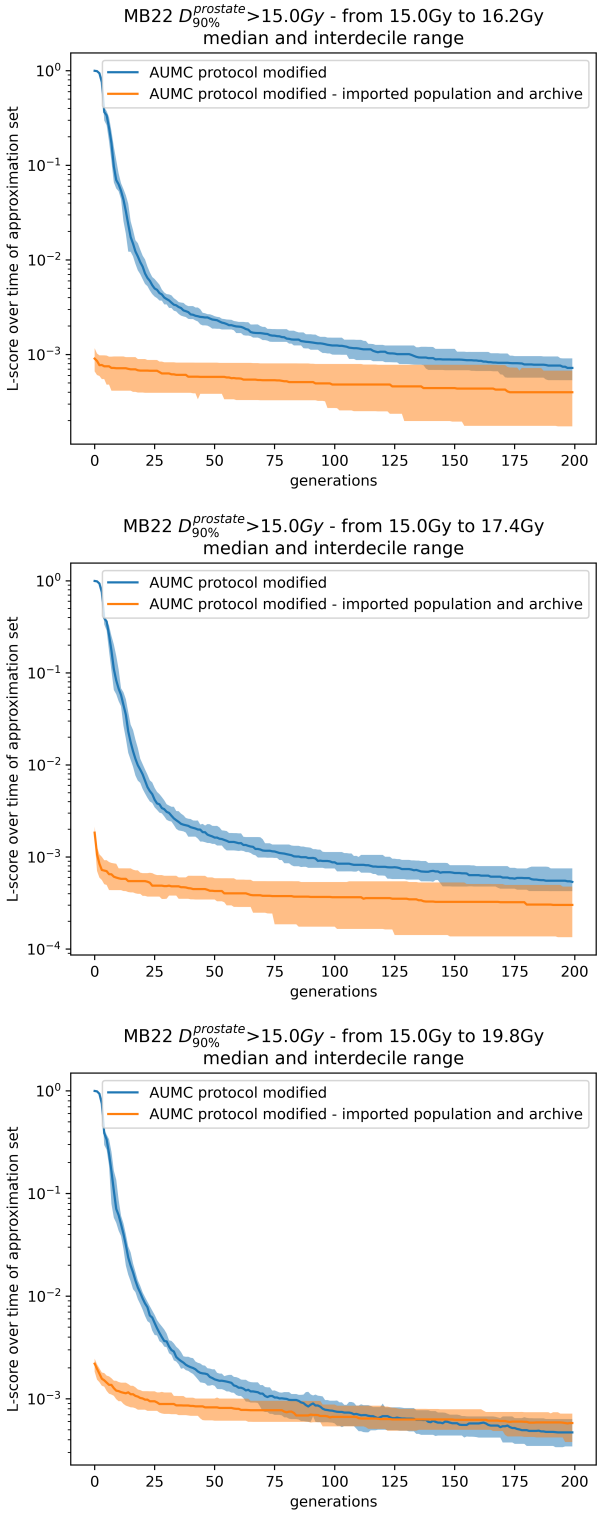


Figure 6.16: Median and interdecile range of the L-score for modifications of 8%, 16% and 32% on aspiration value of $D_{90\%}^{prostate}$ for patient MB22

The advantage is that it does not require initial examples to learn from. However, this method

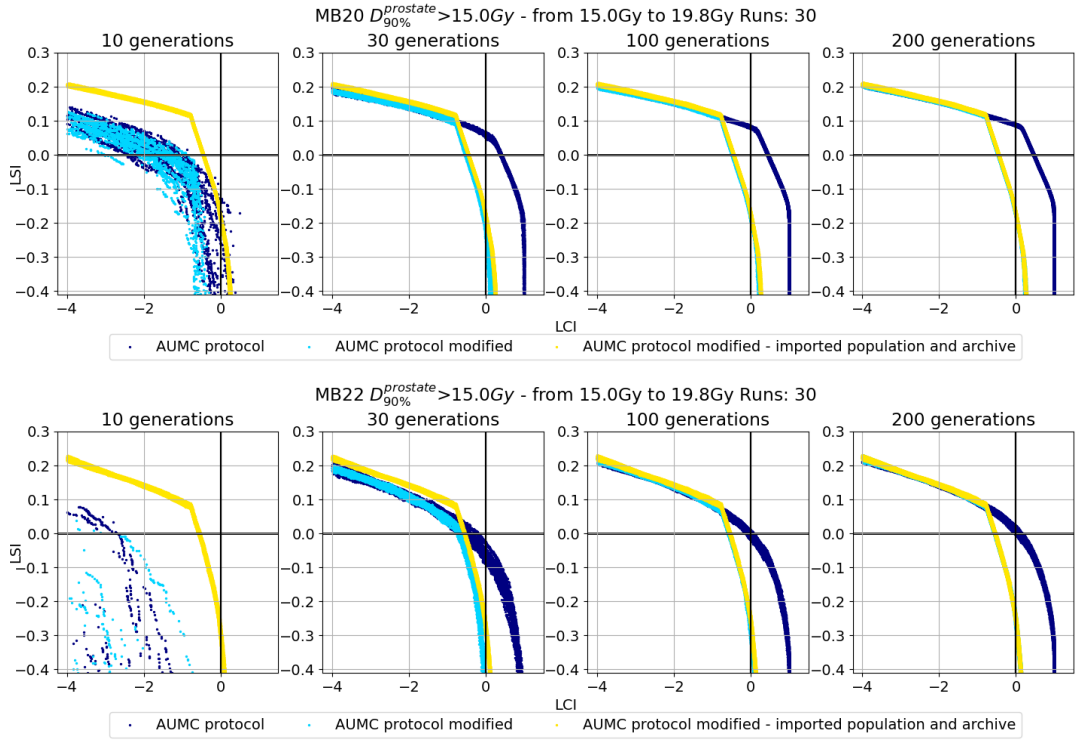


Figure 6.17: LCI/LSI fronts for patients MB20 and MB22 while modifying $D_{90\%}^{prostate}$ with 32% at several generations

still requires the right hyper parameters, such as the range in which dwell positions are affected and the amount by which the dwell times are affected.

The distance metric which will be used is the Euclidean distance between dwell positions and dose points within the organ. So, the lower the Euclidean distance between a dwell position and a dose point within an organ, the more influence the dwell position influences the dose point of the organ. The main difficulty is that each organ has thousands of dose points, meaning it is not trivial to define the influence of a dwell position on an entire organ. In order to get a better idea of the anatomy of a patient and the positions of the dwell positions an overview can be found in Figure 6.18.

As can be seen in Figure ??, the organs have various complex shapes which make it hard to determine the influence of a dwell position on an organ. However, this method does not require exact calculation for for instance influence of a dwell position on an organ, since the goal is to find dwell positions which influence the modified index the most and based on that a modification is made to the dwell times, in order to help boost the convergence time, not to find new (final) solutions.

When a DVI is changed of organ o, all dose points of organ o will be considered. To find the dwell positions which have most influence on organ o, for each dwell position the distance to all dose points is calculated. Based on this distance the inverse square law is used in order to determine the influence, because all organs are considered of the same density and this metric expects, when the distance is twice as large, that the amount of radiation applied will be four times less (the influence is calculated by $\frac{d}{d^2}$ where d is the distance between the dose point and the dwell position). Out of the set of dwell positions the n dwell positions which affect the organ o the most will be selected. These dwell positions are considered to be influencing the DVI which has its aspiration value changed the most. When DVI is a sparing index which has its aspiration value lowered, then the affected dwell positions will have their dwell times lowered and when a

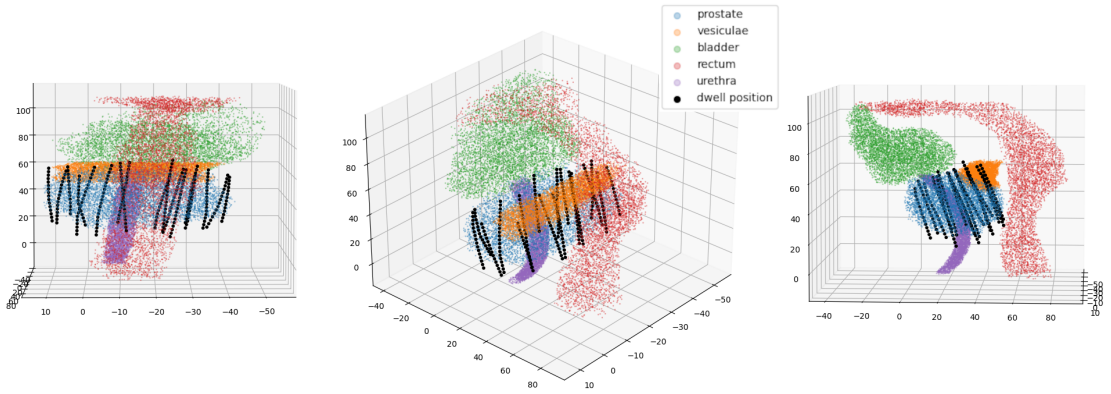


Figure 6.18: This figure shows three different angles of the sampled points of the regions of interest and the locations of the dwell positions for one patient. For the regions of interest 5.000 DCPs are used for visualization purposes

DVI is a coverage index which has its aspiration value increased, then the affected well times will have their dwell times increased. This way the modification of the dwell times will partially compensate for the increase of difficulty of the change of the aspiration value.

When the affected dwell positions have been found the next step is to decide with which factor the dwell times should be changed. This factor should be related to the factor with which the aspiration value of the affected DVI is changed, since a bigger change in a aspiration value would mean that the difference between P^{ρ}_{out} and $P^{\rho'}_{out}$ would also be bigger. Therefore the factor that the dwell times are changed is affected by the amount of change in factor. This factor will be called factor m, the affected dwell positions will have their dwell times changed with m times the change in the aspiration value, so for instance if an aspiration value of a sparing index decreases with 32.0%, when $m = 0.800$, then the decrease in dwell times will be $0.320 * 0.800 = 0.256$, which is by 25.6%.

In order to compensate for the fact that out of the dwell positions with the most influence some of the dwell positions might have more influence, the factor m_w with which each dwell time will be modified will be a weighted version of m. For a dwell position i the weighted factor, m^i_w can be calculated using $m^i_w = \frac{m*A*z^i}{\Sigma^A_{j=0} z^j}$ where A is the total amount of most important dwell positions, z^i the influence of dwell position of i.

The modified plans of the population and the elitist archive will be referred to as $F(P^{\rho}_{out}, \rho')$ and as $F(\mathcal{E}^{\rho}_{out}, \rho')$ respectively.

When modifying dwell positions of plans for sparing indices, dwell times will only be reduced, meaning that when comparing P^{ρ}_{out} and $F(P^{\rho}_{out})$ the modified plans ($F(P^{\rho}_{out})$) will either have a LSI value which is equal or higher and a LCI value which is equal or lower than the unmodified plans (P^{ρ}_{out}). The same holds but in reverse for the coverage indices.

Because of the fact that with this modification solutions can also decrease in quality (by for instance having equal LSI value but worse LCI value when a sparing index is modified), both the modified solutions as well as the original elitist archive will be imported ($\mathcal{E}^{\rho'}_{in} \leftarrow \{F(\mathcal{E}^{\rho}_{out}, \rho'), \mathcal{E}^{\rho}_{out}\}$). Before the first generation of BRIGHT the entire elitist archive is imported and evaluated. After evaluation the Pareto dominated plans will be discarded (because the definition of a plan within the elitist archive is that it should be a non-dominated plan). This means that the initial evaluation of the elitist archive will last twice as long, which is not a substantial amount of time and it does guarantee that the quality of the plans does not decrease when applying the function F.

The result of this modification of the dwell times, for patient MB20, modification of $D^{prostate}_{90\%}$, can be found in Figure 6.19. The conclusion which can be made based on this front is that for the final result, when comparing to rerunning from scratch, that there are barely any differences,

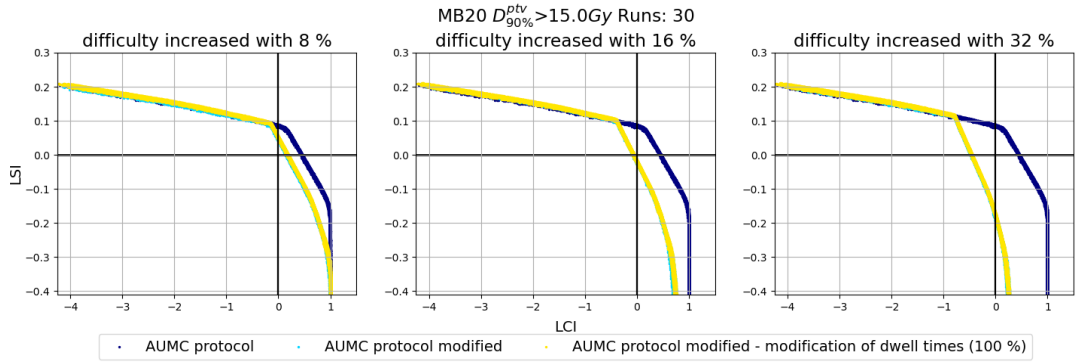


Figure 6.19: Effect on the LCI/LSI fronts after modification of the dwell times for the population and the elitist archive based on a distance metric. These fronts are based on runs which are assumed to be converged (200 generations)

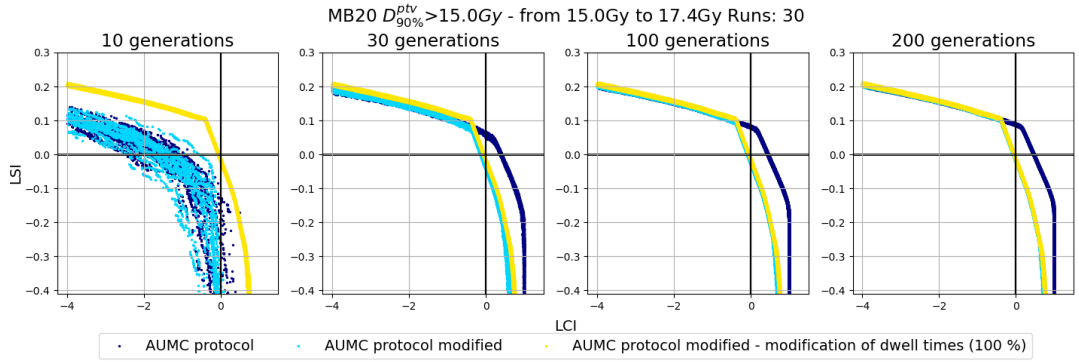


Figure 6.20: Effect on the LCI/LSI fronts after modification of the dwell times for the population and the elitist archive based on a distance metric.

meaning that the modification does not effect the final position of the front.

To see the convergence over time the front will be evaluated at various moments in time, since time is defined in the amount of generations, the front will be shown at different amount of generations. The results for 10, 30, 100 and 200 generations can be found in Figure 6.20. The difference between the imported population and archive and the modification of dwell times with a factor of 100% is hard to see, the reason for this is that those individuals have very similar LCI and LSI, meaning there is barely any difference with regards to convergence.

The lack of improvement is not only visible for this specific change, patient and aspiration value. It seems to be the case for a lot of indices. In most cases modified plans are dominated and therefore instantly discarded from the elitist archive. However, there are some plans which are modified and not dominated, but these modifications often only slightly improve on the current solutions, meaning that especially after a few generations the difference between the original plan and a slightly improved plan is not substantial enough in order to speed up convergence. This is also the reason why other factors of modification are not attempted, since with a low factor the modified plans will not improve sufficiently to make a difference and with a high factor the modified plans will decrease in performance and therefore often being out performed by the non-modified plans.

6.6.2. Plan matching

In Section 5.4 the concept of matching plans in order to find a relation between P^{ρ}_{out} and $P^{\rho'}_{out}$ is explained. In this section the results of plan matching will be shown.

First of all, as also stated in Section 5.4, the goal of plan matching is to find a relation between P_{out}^{ρ} and $P_{out}^{\rho'}$ using similarity between plans to match. If such a relation exists and if such a relation is similar across several patients for a modification on a DVI then this relation can be used in order to modify P_{out}^{ρ} with a function F such that $P_{in}^{\rho'}$ is more similar to $P_{out}^{\rho'}$ than P_{out}^{ρ} and thus decreasing the time required to converge.

In order to find the relation between plans of P_{out}^{ρ} and $P_{out}^{\rho'}$ a similarity metric will be used in order to find the similarity of two plans. The first similarity metric which will be evaluated is a similarity metric based on the unaffected objective, which is either LCI or LSI. When a coverage index is modified, LSI remains unaffected and most differences will be found in the LCI value. This also holds when modifying a sparing index, the LCI value will not change when modifying the protocol and therefore the relation between plans might be able to be tracked by comparing the LSI values. The similarity score between two plans based on the unaffected objective is simply the absolute difference between the values of the unaffected objective (LCI in case of modification of a sparing index and LSI in case of modification of a coverage index).

Another way of finding similarity between plans is by comparing the dwell times of plans. The reason for focusing on the dwell times of the individuals can be found in Section 5.4. The similarity score of two plans will be based on the sum of the absolute differences in dwell times for each dwell position. The lower the similarity score, the more similar the plans are. This means that plans which have the same pattern of dwell times are more likely to have a lower similarity score.

These two metrics will be used in order to create a matrix in which plans from P_{out}^{ρ} and $P_{out}^{\rho'}$ can be compared. The similarity matrix will have all plans of P_{out}^{ρ} on one axis and the plans of $P_{out}^{\rho'}$ on the other axis. Each cell then corresponds to a similarity score between a plan P_{out}^{ρ} and $P_{out}^{\rho'}$.

Examples of these matrices can be seen in Figure 6.21. Within the figure both the metrics for calculating the similarity score as well as two different changes on $D_{90\%}^{prostate}$ can be seen. For both P_{out}^{ρ} and $P_{out}^{\rho'}$ the plans are sorted based on LCI, meaning that it is expected that plans on the diagonal are similar with regards to LSI/LCI, which is also the case as shown in the picture.

When looking at the metric of the absolute LSI differences it can be seen that most similar plans are on the diagonal, which is also expected, since those plans are on similar locations on the front.

However when comparing the two different metrics there is a clear difference, where the absolute differences based on the LSI tends to have more similarities on the diagonal, the sum of absolute differences of the dwell times are mostly similar for plans with low LCI/high LSI. This does make sense since those plans are unlikely to change a lot because the main focus is on the sparing, whereas the change has been made in a coverage index. And at the part of the front where plans have high LCI/low LSI the plans of P_{out}^{ρ} and $P_{out}^{\rho'}$ are less similar. This is also to be expected, since those plans have high LCI, a modification to a coverage index does mean that when plans of P_{out}^{ρ} are evaluated based on ρ' that plans with high LCI will be influenced the most. Meaning that those plans will likely differ the most between P_{out}^{ρ} and $P_{out}^{\rho'}$.

Based on these observations it is most likely that the best similarity metric to use is the one based on the unaffected objective value. However, the metric which is based on dwell times will still be used.

The next step, now the similarity matrix is created, is to further find the relation between the two sets of plans. This is done by matching the plans with the lowest similarity score.

With regards to matching the plans, this will be done one-to-one. Since if the plans are matched one-to-many (from P_{out}^{ρ} to $P_{out}^{\rho'}$) most plans do tend to map to a subset of plans. The reason for this is that for instance the difficulty of a coverage index is increased, then when comparing P_{out}^{ρ} and $P_{out}^{\rho'}$, P_{out}^{ρ} is likely to have lower dwell times in general than $P_{out}^{\rho'}$. In that case plans are more likely to map towards plan which have higher dwell times. This means that two matched plans from P_{out}^{ρ} and $P_{out}^{\rho'}$ will have less difference since most of the plans will find a close match on another place on the front. This is also the reason for not using many-to-one, since when for

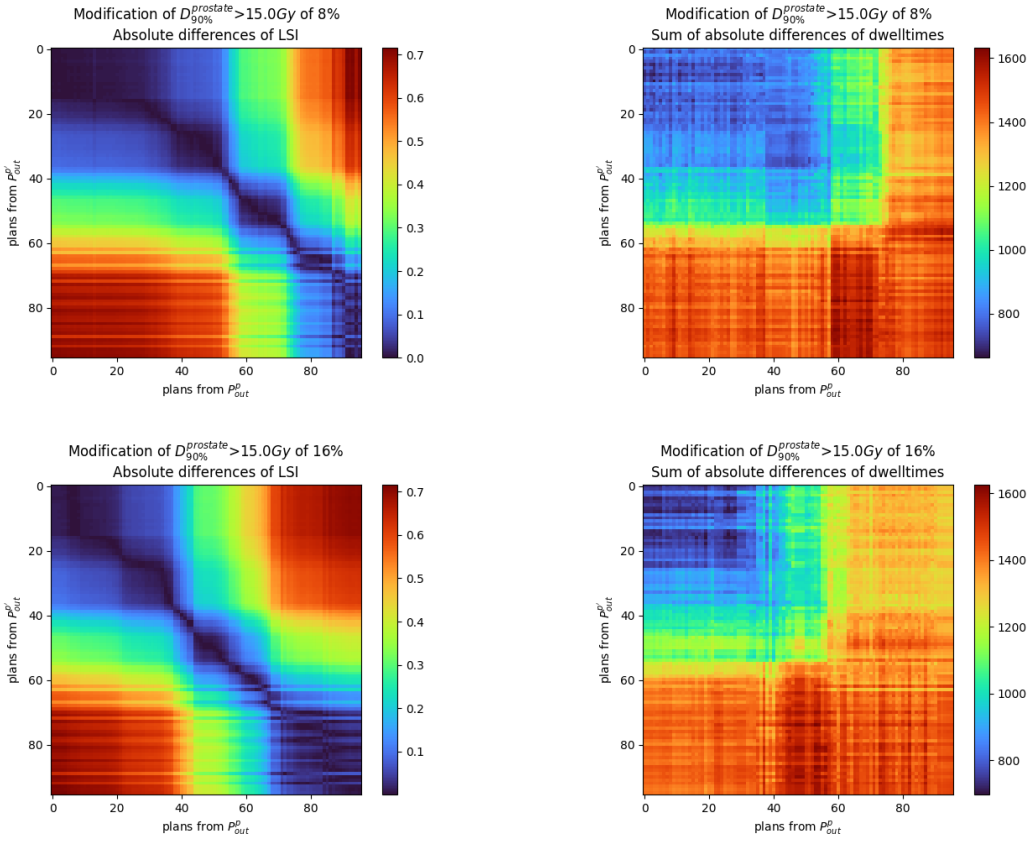


Figure 6.21: Similarity scores for patient MB20 when comparing P_{out}^{ρ} and $P_{out}^{\rho'}$. The lower the similarity score the more similar the plans are.

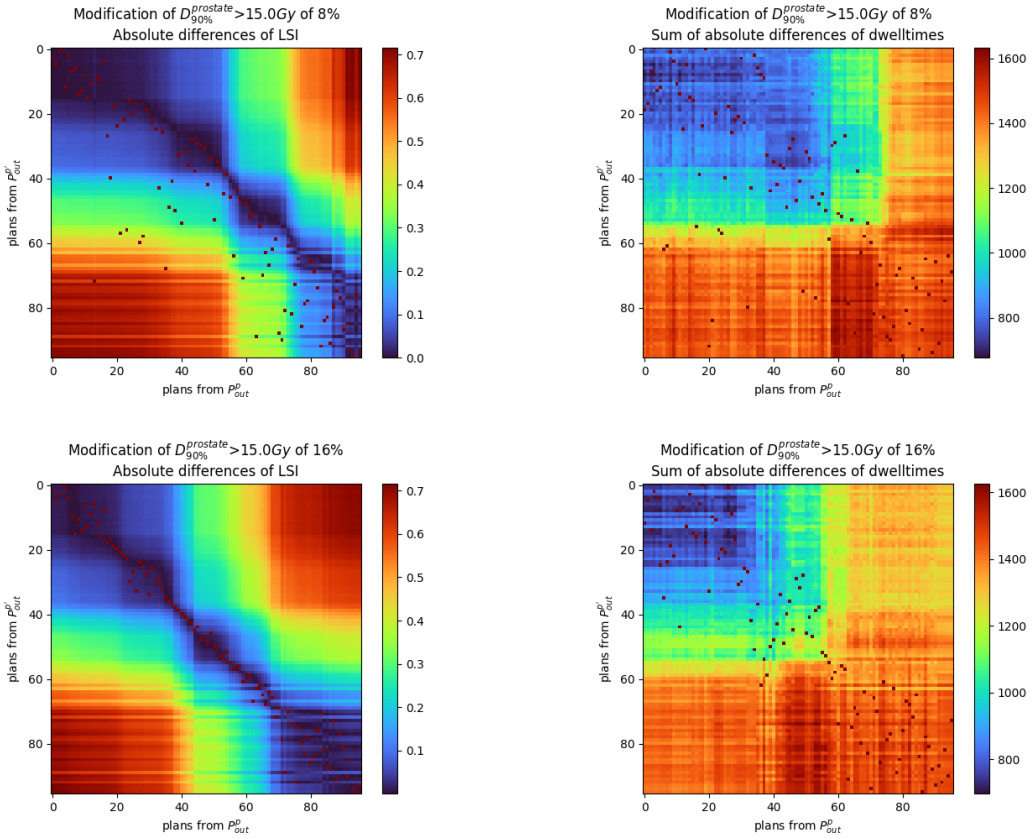


Figure 6.22: Similarity scores for patient MB20 when comparing P_{out}^{ρ} and $P_{out}^{\rho'}$. The dark red dots indicate a mapping between the plans

instance a coverage index the aspiration value is increased, plans with relatively low coverage in P_{out}^{ρ} (and therefore lower dwell times) will not be matched due to plans in $P_{out}^{\rho'}$ having higher values for the dwell times, making the plans with high coverage in P_{out}^{ρ} too influential.

For this reason the method used to map the plans will be one-to-one. The mapping will prioritize to have the sum of similarity scores of matched plan to be as low (as similar) as possible. The mapping for the example case can be found in Figure 6.22. When a point has a dark red color it means that those plans from P_{out}^{ρ} and $P_{out}^{\rho'}$ are matched.

As can be seen there is a slight trend which can be seen across the four subfigures, in general the matched plans tend to be across the diagonal, meaning there is a tendency to match the plans which have a similar index. The plans are sorted based on their LCI values, since there is a trade-off between LCI and LSI this also means that in almost any case the plans are also sorted on LSI (but with the inverse sorting method compared to LCI). However, there is a clear difference between matching based on LSI and matching based on the sum of absolute differences in dwell times. As the similarity matrices already did show, the similarity metric of the LSI differences shows a pattern which is much more clear. The matches of the absolute differences in dwell times also seem to be more scattered.

Another interesting conclusion which could be drawn from Figure 6.22, namely that in this specific case it seems that a modification of 16% of $D_{90\%}^{prostate}$ seems to lead to less spread around the diagonal than a change of 8%. This could have multiple reasons, but likely the main reason for this is that the distribution of plans on the front is more similar.

These mappings can also be visualized in a LCI/LSI front way, with the mapping between two plans represented as a line. The results of these mappings showing the relation based on the

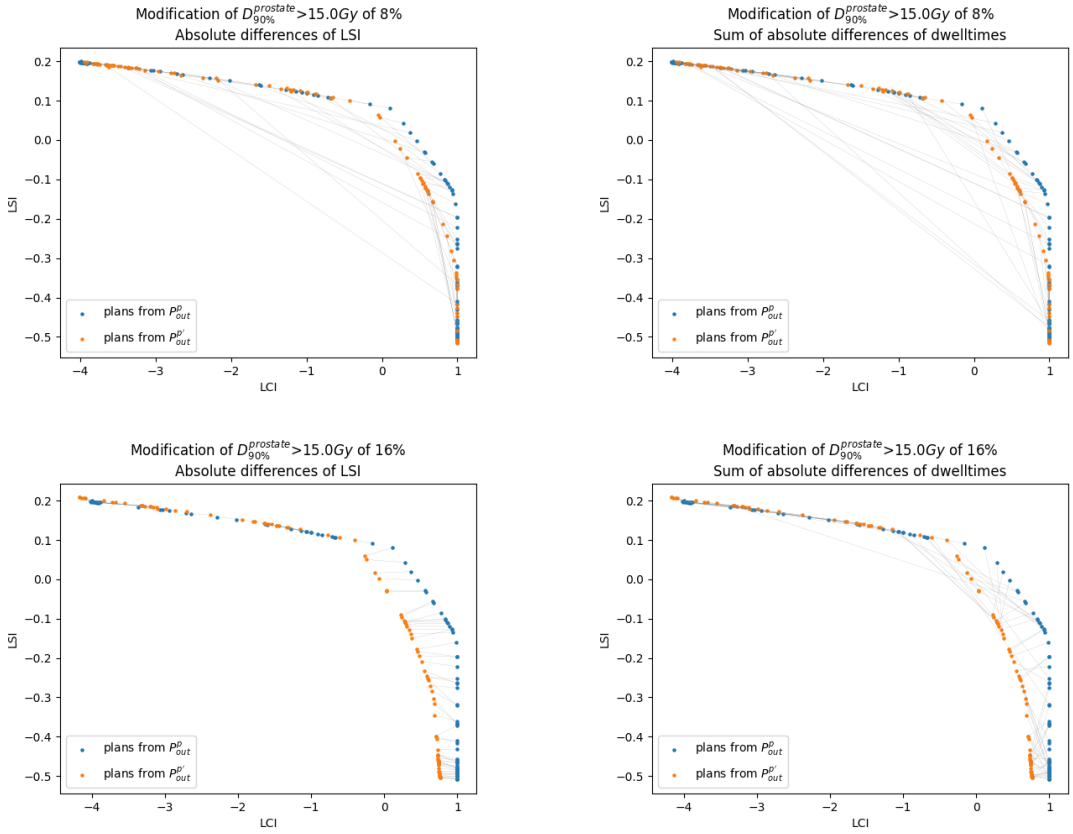


Figure 6.23: LCI/LSI fronts for MB20 with plan matching when modifying $D_{90\%}^{prostate}$

LCI/LSI front can be seen in Figure 6.23. These mappings are the same as in Figure 6.22, but visualised in a different way. As expected the results also do look similar to the similarity matrix mappings.

One disadvantage of using a one to one mapping based on these results is that the mapping also seems to dependent on the distribution of the individuals on the front. When looking at the difference between a change of 8% and 16% with the LCI differences as similarity metric, the mapping of the change of 8% seems to have less of a pattern in the matches of the plans, where as the matches of the plans with a change of 16% seem to have a mapping in which there is a clear pattern visible, based on the LSI scores, on which the similarity metric is based.

Based on these matched plans the average change of dwell times can be calculated (for example, based on two runs, if plan with index 0 is matched once with a plan with index 2, and once with index 4, the differences in dwell times between run 1, plan 0 and plan 2 and the differences in dwell times between run 2, plan 0 and plan 4 will be averaged). Using these average changes in dwell times a relation between P_{out}^{ρ} and $P_{out}^{\rho'}$ could be found.

The comparison of matched plans for the modifications of the aspiration value of $D_{90\%}^{prostate}$ can be found in Figure 6.24.

As can be seen at some points the matched individuals do have similar differences when comparing to plans that have similar LCI/LSI values (which have a similar index value). This means that for those dwell positions from P_{out}^{ρ} to $P_{out}^{\rho'}$ the dwell times will generally either increase or decrease in dwell times on average, which indicates a pattern.

Another interesting observation is that there are a few dwell positions which have their changes in dwell times similar for the first 30-40 indices of the population. These plans are the plans with low LCI/high LSI, meaning at the top left corner of the LCI/LSI figures. As can be seen, for those

30-40 plans the LCI and LSI values for P_{out}^{ρ} and $P_{out}^{\rho'}$ are similar here. These first 30-40 indices are mainly the indices which are affected less by the modification of $D_{90\%}^{prostate}$ (this can also be confirmed when looking at the worst performing coverage indices LCI/LSI front in Section 6.3). This does however not mean that the plans do not change, even though it could also indicate a sub optimal mapping.

For the other plans (with an index higher than 30-40) the mapping seems to have less of a pattern, even though there are small groups for some dwell positions with similar changes, there is no clear pattern.

When comparing the two different metrics it is clear that the matching, even though they are not really similar in which plans they map, the differences in dwell times for the matched plans do show similarities at some points. The aforementioned pattern for the plans with relatively high LSI/relatively low LCI can be seen for both the plan matching sets based on the two different similarity metrics.

The difference between a modification of 8% and 16% is more substantial than the differences between the two different similarity metrics. Even though the unaffected dwell positions remain similar, for most of the dwell positions it is hard to find a relation between a modification of 8% compared to a modification of 16%.

When a relation between P_{out}^{ρ} and $P_{out}^{\rho'}$ based on the modification is found, one of the parameters should be the amount the aspiration value is modified. Therefore a change of 32% of $D_{90\%}^{prostate}$ for patient MB20 is also considered. The results can be found in Figure 6.25.

When comparing the changes of 8% and 16% with the change of 32% the same conclusion could be made as when 8% and 16% were compared, the unaffected dwell positions remain mostly the same, however, the dwell positions which are changed between P_{out}^{ρ} and $P_{out}^{\rho'}$ do show some similarities, but are for most dwell positions different.

These differences in dwell times between the plans which are matched are then used as input for function F. The modification which F applies is based on the average difference in dwell times based on the matched plans.

With some initial experiments using the mapping found it was clear that $P_{in}^{\rho'} \leftarrow F(P_{out}^{\rho}, \rho')$ did not give any substantial improvements in convergence. The lack of substantial improvements was still there when the elitist archive was imported and modified ($P_{in}^{\rho'} \leftarrow F(P_{out}^{\rho}, \rho')$).

In order for this method to speed up the convergence, the plan matching could not be calculated on the spot, since that would require $P_{out}^{\rho'}$ to be known. But $P_{out}^{\rho'}$ is the goal (for which a faster convergence is sought after), so this would require a model which is pre-trained. Therefore the figures for a change in $D_{90\%}^{prostate}$ for patient MB20 should be compared to the results of plan matching based on a different patient.

The results in plan matching for patient MB21 can be found in Figure 6.26. Again, as also the case with patient MB20, it shows some dwell positions for which plans with similar indices have similar modifications from P_{out}^{ρ} to $P_{out}^{\rho'}$.

However, when comparing Figure 6.24 and Figure 6.26 (the plan matching differences in dwell times for patient MB20 and patient MB21), it can be seen that both have completely different changes at different dwell positions. The main reason for this is likely that first of all the locations of the dwell positions between the patients do differ, but also the anatomy between the patients is different. The relation between P_{out}^{ρ} to $P_{out}^{\rho'}$ is very patient dependent and therefore it is hard to find a uniform mapping, even though when the goal is not to find a function for which holds $F(P_{out}^{\rho}, \rho') = P_{out}^{\rho'}$ but only to find a way to modify P_{out}^{ρ} to be more similar to $P_{out}^{\rho'}$ it is hard to find a mapping which fits for all patients. Another reason for the differences could be that the mapping is not optimal and that there might be other ways to map the plans. Possible solutions for this will be discussed in Chapter 7.1.

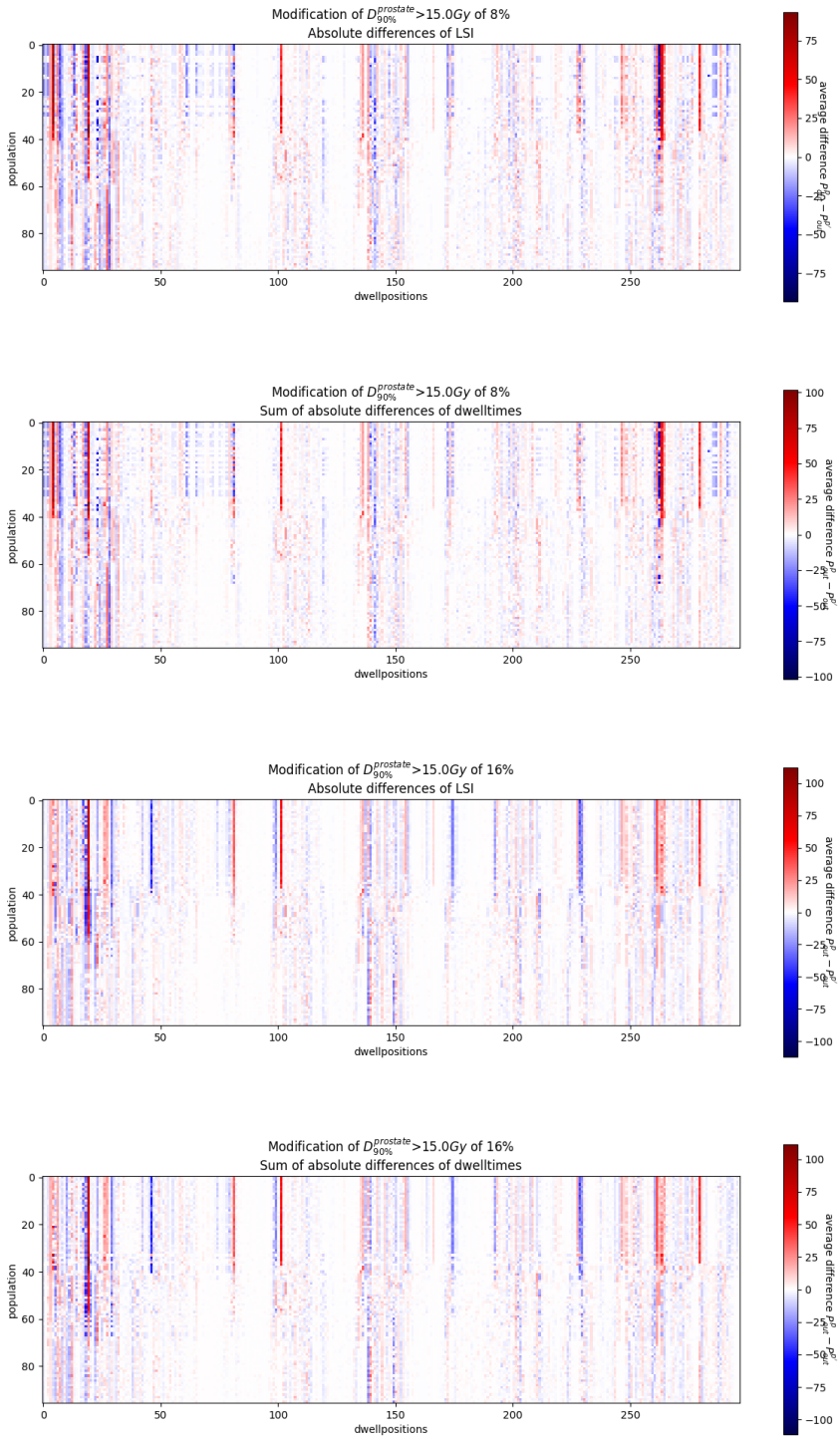


Figure 6.24: Average differences for plan matching of patient MB20 when modifying $D_{90\%}^{prostate}$

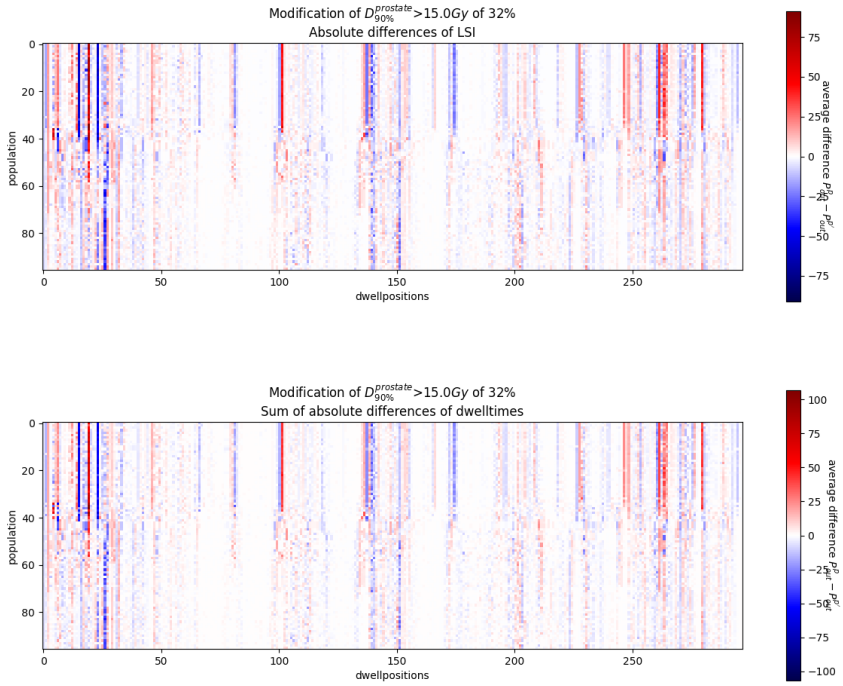


Figure 6.25: Average differences for plan matching of patient MB20 when modifying $D_{90\%}^{prostate}$

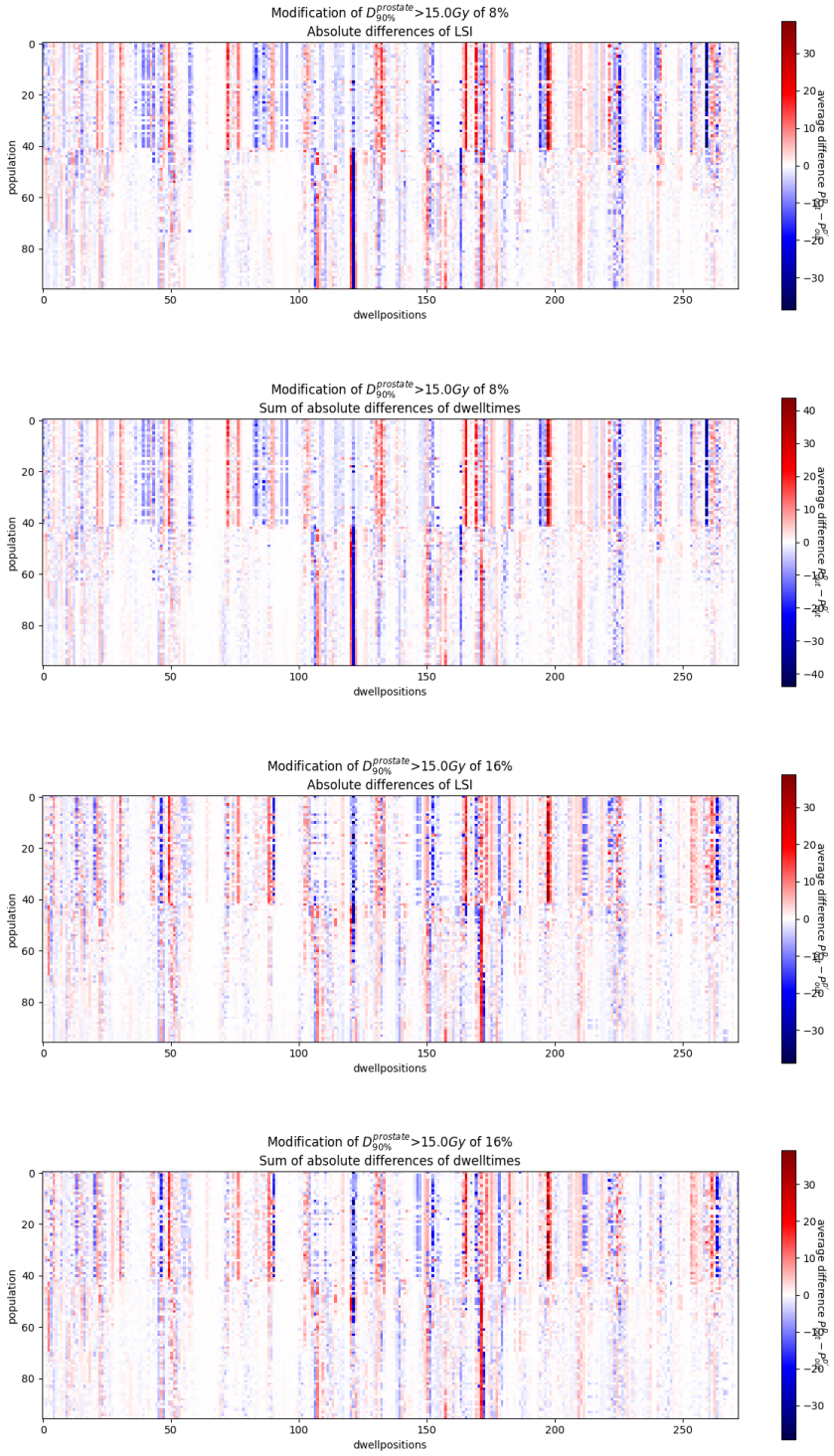


Figure 6.26: Average differences for plan matching of patient MB21 when modifying $D_{90\%}^{prostate}$

7

Discussion and conclusions

The main goal of this research was to find a way to decrease time required to converge to a set of high-quality plans once the protocol has been changed. The methods used were not previously used and provide a way in order to improve the time required to converge. Besides achieving the goal of lowering the time required to converge this research also gives additional insights into the effects when changing one of the DVIs. Also a way to analyse whether a change of the protocol has effect on the plans has been presented.

In the upcoming sections further possible research suggestions will be given, as well as possible short-comings of this research. In the last section the conclusion of this research will be presented.

7.1. Discussion

The described method of importing the population and the elitist archive does allow to decrease the time required to converge to a set of high-quality treatment plans after modification of the protocol. This improvement would likely be useful in clinical practice. However, there are possibly still chances in order to further improve the time required. Even though the plan matching did not provide improvement in convergence speed, there still might be ways in which the plan matching could be used in order to improve the speed of the convergence. There would likely be two major modifications required to use plan matching to find a good model which can be used to modify the plans to be closer to the expected final result. The first is that since it would have to work on multiple patients with different catheter placements (and therefore, different dwell locations, different number of catheters and different anatomy) that the index of a dwell position is not sufficient. Instead, a general location would have to be used for the location of the dwell position in relation to the organs. Secondly, when the first modification is done, an experiment using a lot of patients should be executed in order to see whether there is a clear pattern. A way of doing this (for one modification at the time) would be to optimize for a large set of patients (i.e. 20+ patients) on the AUMC protocol. Then, the same should be done again but with the modification applied to the AUMC protocol. Both sets should have a sufficient amount of runs to deal with the uncertainty. After obtaining the results some kind of pattern analysis between the runs on the original AUMC protocol and the runs on the modified AUMC protocol could be performed, this could also confirm whether there even is a pattern which can be exploited in order to decrease the time required to converge. Another method which could be used is some kind of supervised learning, the subject which will be 'trained' will be a model for modifying the dwell times, the input will be a set of individuals optimized on the original AUMC protocol, the output will be a set of individuals with their dwell times modified. The fitness of the output will be defined based on the amount of generations which are required to converge (various definitions of convergence can be used here, for instance, a value-to-reach for LCI and LSI or lack of improvement over a few generations). However, the fitness of the output could not directly be used to tweak the modification model, since it does not contain any information on whether the modification should

be increased or decreased. For that additional information such as the progression of the dwell times over the generations within the elitist archive can be used, that way it could be tracked which dwell times do increase or decrease. This trend in dwell times can then be used to tweak the model. The magnitude of the modification can then be determined by the fitness score (i.e., for relatively low fitness the modification would be larger than for high fitness).

Another possible improvement when considering a step in which the dwell times are modified based on a comparison of a run on the original protocol and a run on the modified protocol is to rely less on the index of dwell positions, but more on their location within the body. The main reason for doing this would be that it would be easier to find a pattern for modifying the dwell times across multiple patients, since using the indices of the dwell positions makes the approach patient dependant. A possible way of doing this is by defining a dwell position (for comparison purposes) by its distances to each of the organs.

A possible addition to this research could be the addition of DTMR (Dwell Time Modulation Reduction), which forces a smoother distribution of dwell times. This could have an effect on the effect of modifying the protocol but also on the main way to improve convergence after a change in the protocol, which is importing the population and the elitist archive. Smoother distribution of the dwell times will possibly mean that it is harder to shift radiation whenever the protocol is changed, which happens when optimization is done using imported population and elitist archive. However, the effect of DTMR on plan matching could be positive since plans are most likely better to match based on characteristics when using DTMR. The reason why it could have a positive effect is that the dwell time distribution is more smooth (less difference between dwell positions which are within close proximity of each other), meaning that plans before and after the modification of the protocol are likely to be more similar than without using DTMR. If plans are more similar before and after the modification it is to be expected that convergence will take less time.

Another point we have considered is the fact that the protocol that is used and what would happen if another protocol is used. It is likely that all the conclusions and concepts from this research also apply to another protocol, but in some cases, differences might be more or less substantial. The effect of adding a GTV (gross tumor volume) coverage index would also likely change the effect a modification to the protocol has since the GTV will be smaller than the PTV.

Lastly, in order to be able to really answer the research question with regards to clinical practice, this study should be expanded to a larger number of patients, and an observer study would be required. In this observer study the clinical relevance of the improvements will be assesed by experts. The difference between two sets of fronts is hard to quantify with clinical relevance. While importing the population and the elitist archive, the final (converged) front is found faster. However, it is not possible to say how clinically significant the advantage of importing the population and elitist archive is without any clinical expertise.

7.2. Conclusion

In order to generate high-quality treatment plans for high-dose-rate brachytherapy, the evolutionary algorithm MO-RV-GOMEA can be used. This application of MO-RV-GOMEA on HDR brachytherapy is called BRIGHT. These sets of high-quality treatment plans do trade-off between giving more dose to the target regions and sparing organs at risk more.

The main goal of this research was to evaluate the effect of changes in the protocol and to improve convergence speed towards high-quality treatment plans when the protocol is changed.

The first conclusion which could be made about the effect of changes within the protocol is that some aspiration values of DVIs can be modified without having a substantial effect on the LCI/LSI values of the front. For this research, the values for which the aspiration values of the DVIs, as specified in the clinical protocol, can be modified without having a substantial impact on the LCI and LSI, is called the slack value of the DVI. The slack value is patient dependent, and which values have slack differs from patient to patient. The slack value is calculated based on a set of plans. When the slack value is calculated based on the entire front, the slack will be less, and in most cases, there is barely any slack when considering the entire front. When only

the set of plans within the golden corner are considered (which are the plans which are the most interesting for an expert since it does satisfy the protocol), there is more slack. For example, $V_{150\%}^{prostate}$ and $V_{200\%}^{prostate}$ seem to have a substantial amount of slack when only considering plans within the golden corner. The calculation of the slack does not require a lot of run time and could be used in order to decide not to re-run BRIGHT because the results would end up similar to the initial run.

Secondly, the effect of changes within the protocol is evaluated. Only one change of an aspiration value of one of the DVIs is considered at the time in order to isolate the effect of such changes. As expected, when changing an aspiration value value to make a DVI harder to achieve, the LCI/LSI value is influenced when the change is outside of the slack range. Plans for which the DVI which is changed that constitutes the most to the worst-performing index in their category (coverage or sparing), are more influenced by a change. In some cases, similar LCI/LSI fronts can be achieved when compared with a run on the original protocol, but in most cases (for changes that increase the difficulty by 8% or more), the front will either have worse LCI or worse LSI compared to the run on the original protocol. Another conclusion is that based on the experiments is that the effect of a change is patient dependent, meaning that a change in a specific DVI will have other effects on the results based on which patients run the change is applied.

In order to achieve a front of high-quality treatment plans without the need to run BRIGHT from scratch, one investigated method was to use existing information that was based on the initial run (a so-called 'warm-start'). It was shown that importing the population and the elitist archive improved the convergence speed of BRIGHT following the modification of the clinical protocol, compared to running BRIGHT from scratch. The population is usually randomly initialized, and the elitist archive is in a default case based on the randomly initialized population. Even though the plans of the initial run are not optimized towards the protocol with the modification, it still helps with convergence. The main risk identified with the approach of importing the population and the elitist archive is that it could result in BRIGHT becoming stuck in a local optimum. This is evaluated and when importing the population and the elitist archive, this is not the case, as indicated by the fact that, after 200 generations, both the run from scratch on the modified protocol as the run with imported population and elitist archive, achieve fronts of near-identical quality. The amount of improvement in convergence time is hard to quantify since it is hard to compare two fronts based on clinical relevance. However, when a clinical expert is exploring, the direction of the improvement of the plans can be found after significantly fewer generations when importing the population and the elitist archive. The amount of time (generations) saved depends on a lot of factors, for instance, which DVI is modified and with what amount. For example, a modification on a DVI that has no or little slack (often, for instance, $D_{90\%}^{prostate}$) will require more time than a similar modification on a DVI that has a substantial amount of slack (often, for instance, $V_{200\%}^{prostate}$). However, in general, when importing the population and the elitist archive plans, which are similar to the final plans (which are found after 200 generations) can be found within as little as 5-10 generations, whereas this is in about 30 generations for random initialization. The definition of convergence to high-quality treatment plans is hard to define, but based on the LCI/LSI fronts over generations, when using the imported population and elitist archive, the plans will not change substantially after 30 generations, whereas with random initialization, the plans do still change up until at least 100 generations. The time saved in seconds when comparing importing the population and the elitist archive is also highly dependent on the setup which is used. When calculating the time saved based on the setup (approximately 0.8 seconds per generation), the time saved on convergence is 56 $(70 * 0.8)$ seconds approximately.

In order to try and improve the convergence time, even more, a modification attempt has been made based on the imported population and elitist archive. The general idea was to modify the dwell times of the plans which were being imported in order to make these plans better by taking into account the change in the DVI. The first modification which was done was based on the distance to the organ for which the aspiration value was changed. However, this did not improve convergence time, as the convergence speed was similar to just importing the population and elitist archive without any modifications. The likely reason for this was that the methodology

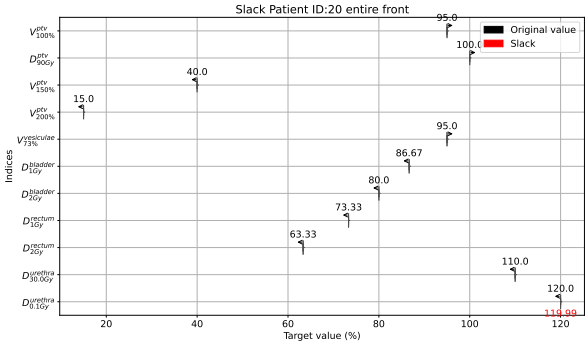
used was too simple and that most modified plans would be Pareto dominated by the original plans. Furthermore, when an improvement was found, such an improvement was not substantial enough to positively influence the convergence time.

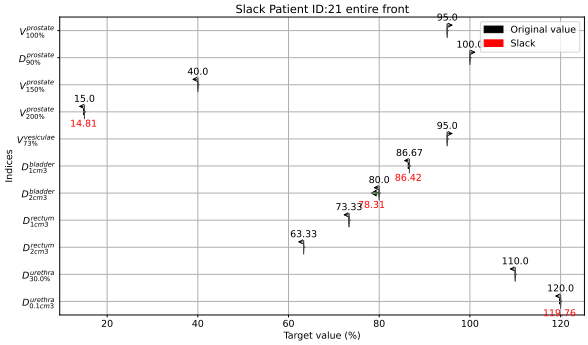
Another method of determining the modification in dwell times is to match plans from before and after the protocol change in order to find a mapping to modify the dwell times. The matching of the plans did show promising results, as there was also a pattern in the plans before and after the change of protocol, however, these patterns were completely different from patient to patient. It also did not show significant improvements in the convergence time. The reason for this is possibly that the plan matching was not optimal and that trying to track plans from before and after a protocol change is unfeasible since plans change a lot.

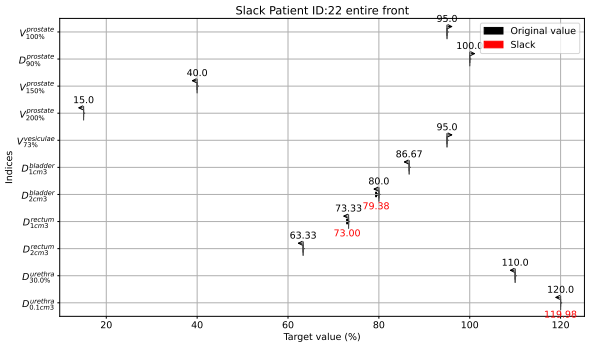
The main contribution is the decrease in time required to converge to high-quality treatment plans after modification of the protocol after an initial run. This is achieved by importing the population and the elitist archive of an initial run in order to improve the time required to converge to a set of high-quality treatment plans. The increase in performance can reach up to 50-70 generations (approximately 40-60 seconds on the setup used in this research), dependent on which DVI is modified and by how much the DVI is modified. Furthermore, this research also provides insight into the effects of modifications to the protocol, which can be used to provide more insight to the clinical experts. The slack can be used in order to decide whether re-optimization is required after a modification within the protocol and also gives further insight on the possible trade-offs between the coverage and sparing aspiration values.

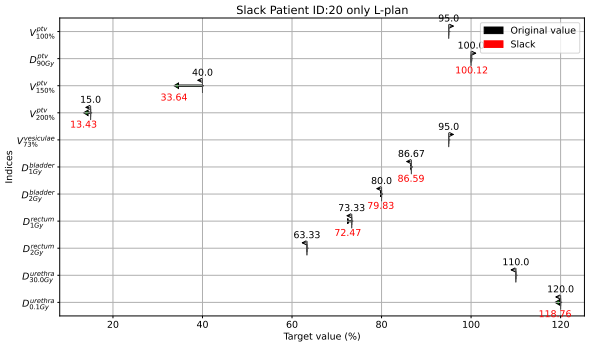
8

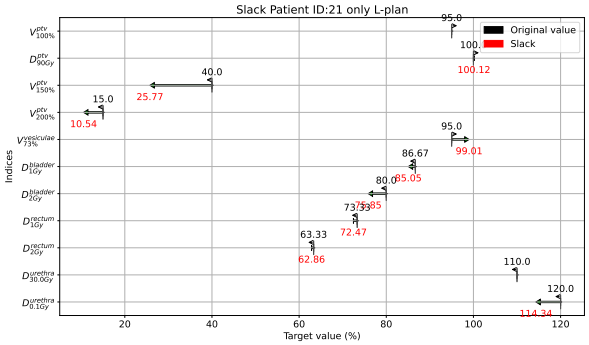
Appendix

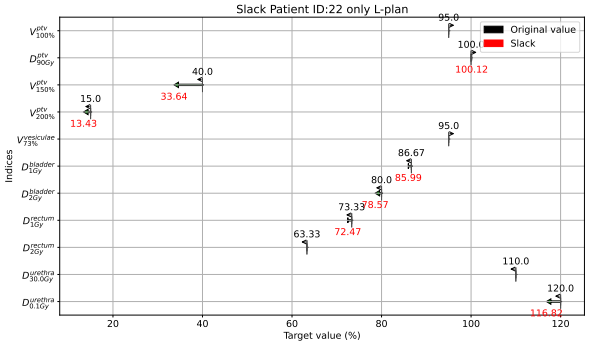

Slack Patient ID:20 entire front
Original value
Slack
Indices
Target value (%)
95.0
40.0
15.0
95.0
86.67
80.0
73.33
63.33
110.0
120.0
119.99

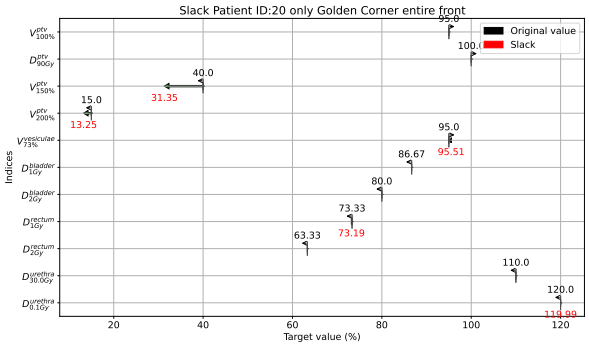

Slack Patient ID:21 entire front
Original value
Slack
Indices
Target value (%)
95.0
40.0
15.0
14.81
95.0
86.67
80.0
86.42
78.31
73.33
63.33
110.0
120.0
119.76

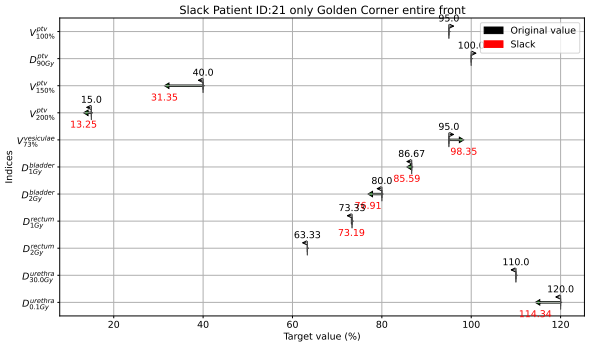

Slack Patient ID:22 entire front
Original value
Slack
Indices
Target value (%)
95.0
100.0
40.0
15.0
95.0
86.67
80.0
79.38
73.33
73.00
63.33
110.0
120.0
119.98

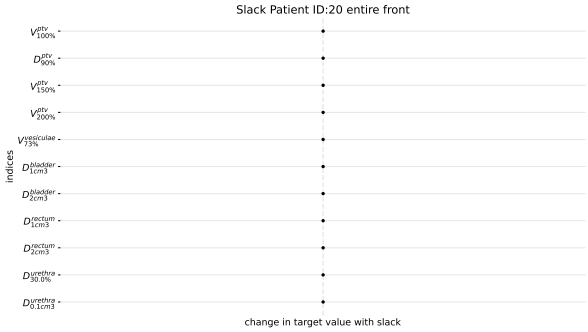

Slack Patient ID:20 only L-plan
Original value
Slack
Indices
Target value (%)
95.0
100.12
40.0
33.64
15.0
13.43
95.0
86.67
86.59
80.0
79.83
73.33
72.47
63.33
110.0
120.0
118.76

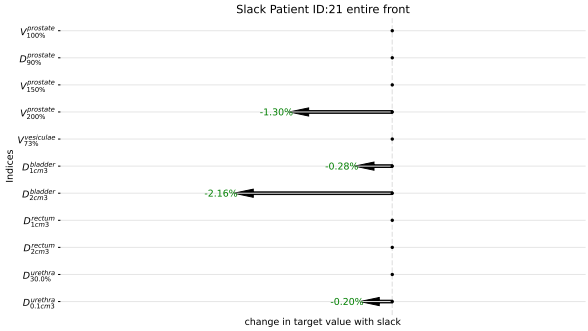

Slack Patient ID:21 only L-plan
Original value
Slack
Indices
Target value (%)
95.0
100.12
40.0
25.77
15.0
10.54
95.0
99.01
86.67
85.05
80.0
73.33
72.47
63.33
62.86
110.0
120.0
114.34

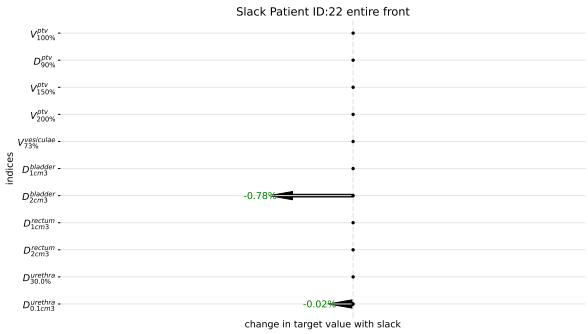

Slack Patient ID:22 only L-plan
Original value
Slack
Indices
Target value (%)
95.0
100.12
40.0
33.64
15.0
13.43
95.0
86.67
85.99
80.0
78.57
73.33
72.47
63.33
110.0
120.0
116.82

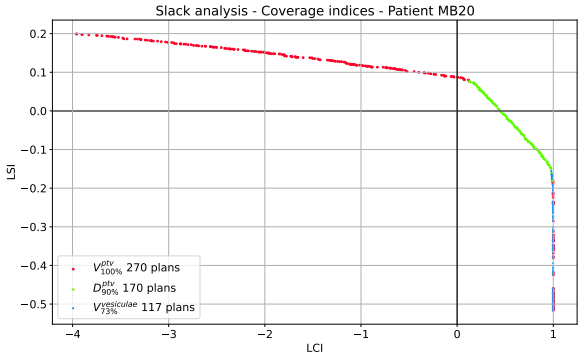

Slack Patient ID:20 only Golden Corner entire front
Original value
Slack
Indices
Target value (%)
95.0
100.0
40.0
31.35
15.0
13.25
95.0
95.51
86.67
80.0
73.33
73.19
63.33
110.0
120.0
119.99

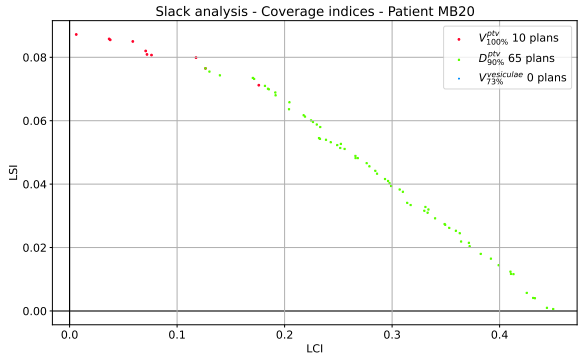

Slack Patient ID:21 only Golden Corner entire front
Original value
Slack
Indices
Target value (%)
95.0
100.0
40.0
31.35
15.0
13.25
95.0
98.35
86.67
85.59
80.0
75.91
73.33
73.19
63.33
110.0
120.0
114.34

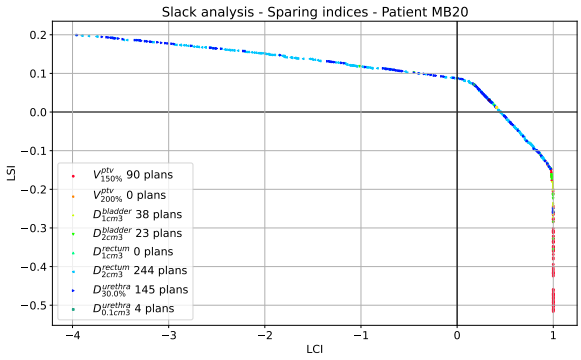

Slack Patient ID:20 entire front
indices
change in target value with slack

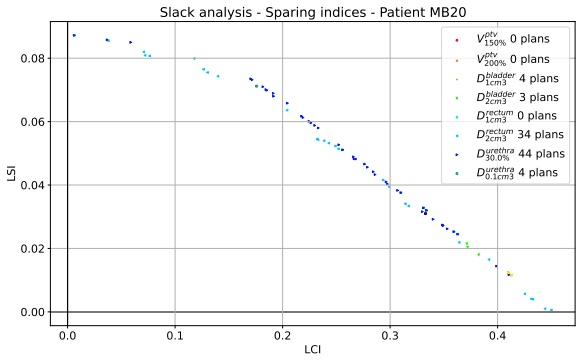

Slack Patient ID:21 entire front
indices
-1.30%
-0.28%
-2.16%
-0.20%
change in target value with slack

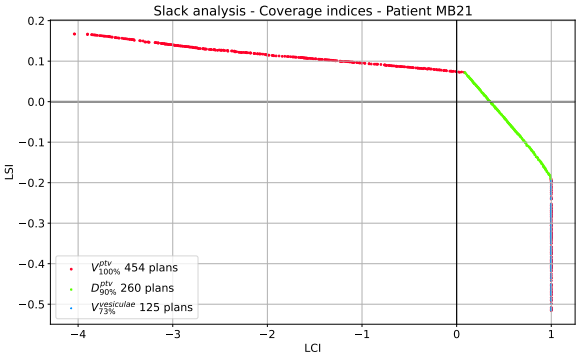

Slack Patient ID:22 entire front
indices
-0.78%
-0.02%
change in target value with slack

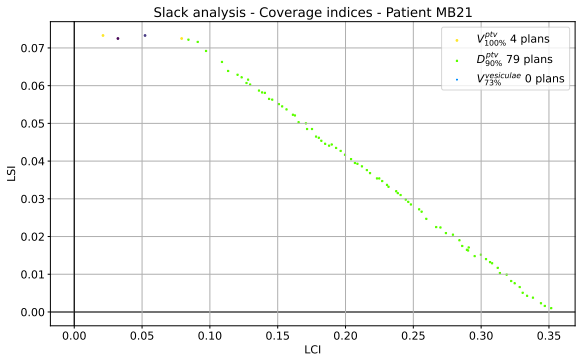

Slack analysis - Coverage indices - Patient MB20
LSI
LCI
270 plans
170 plans
117 plans

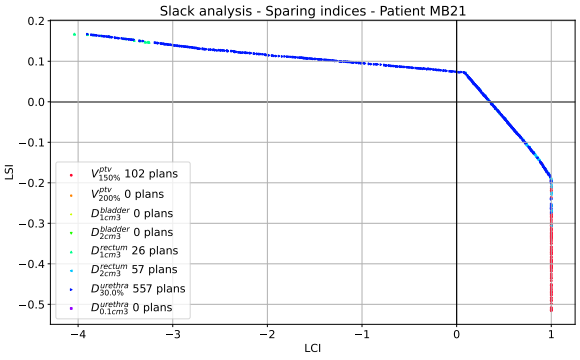

Slack analysis - Coverage indices - Patient MB20
LSI
LCI
10 plans
65 plans
0 plans

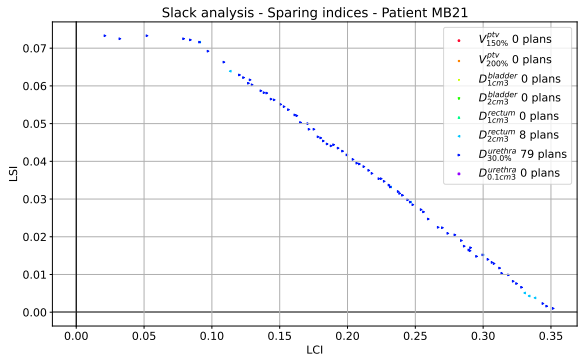

Slack analysis - Sparing indices - Patient MB20
LSI
LCI
90 plans
0 plans
38 plans
23 plans
0 plans
244 plans
145 plans
4 plans

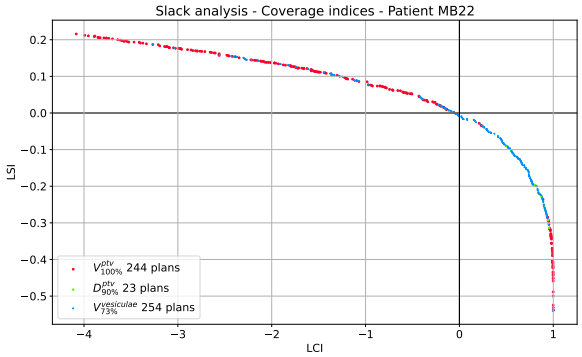

Slack analysis - Sparing indices - Patient MB20
LSI
LCI
$V^{ptv}_{150\%}$ 0 plans
$V^{ptv}_{200\%}$ 0 plans
$D^{bladder}_{1cm3}$ 4 plans
$D^{bladder}_{2cm3}$ 3 plans
D^{rectum}_{1cm3} 0 plans
D^{rectum}_{2cm3} 34 plans
$D^{urethra}_{30.0\%}$ 44 plans
$D^{urethra}_{0.1cm3}$ 4 plans

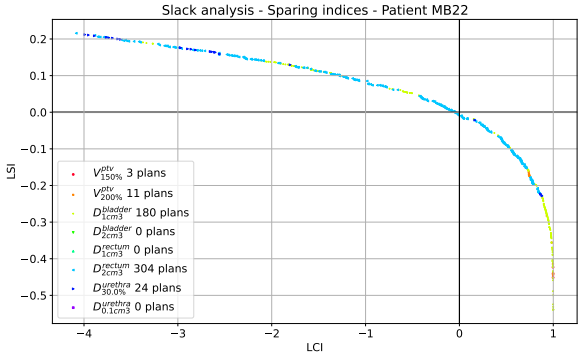

Slack analysis - Coverage indices - Patient MB21
LSI
LCI
$V^{ptv}_{100\%}$ 454 plans
$D^{ptv}_{90\%}$ 260 plans
$V^{vesiculae}_{73\%}$ 125 plans

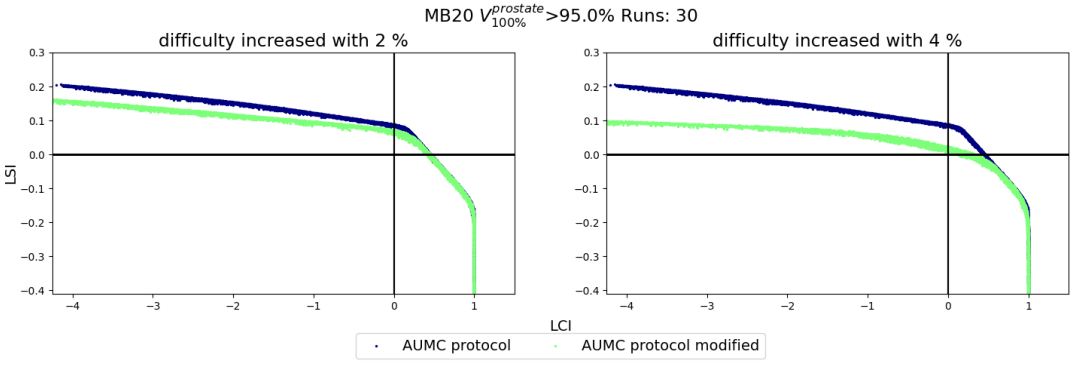

Slack analysis - Coverage indices - Patient MB21
LSI
LCI
$V^{ptv}_{100\%}$ 4 plans
$D^{ptv}_{90\%}$ 79 plans
$V^{vesiculae}_{73\%}$ 0 plans

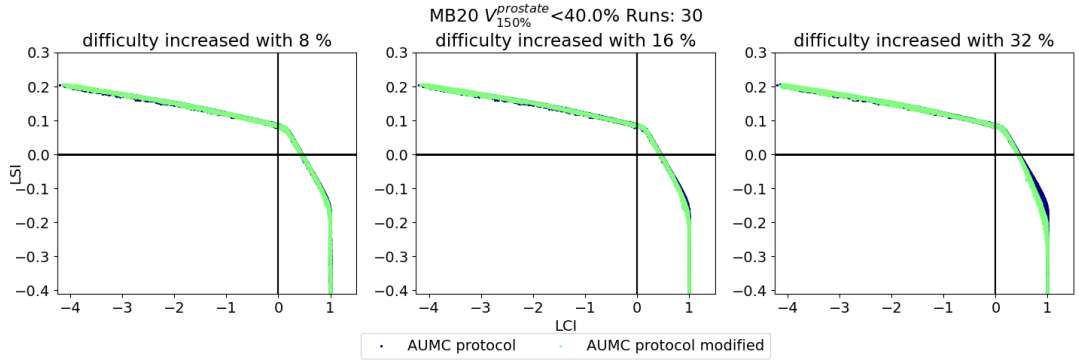

Slack analysis - Sparing indices - Patient MB21
LSI
LCI
$V^{ptv}_{150\%}$ 102 plans
$V^{ptv}_{200\%}$ 0 plans
$D^{bladder}_{1cm3}$ 0 plans
$D^{bladder}_{2cm3}$ 0 plans
D^{rectum}_{1cm3} 26 plans
D^{rectum}_{2cm3} 57 plans
$D^{urethra}_{30.0\%}$ 557 plans
$D^{urethra}_{0.1cm3}$ 0 plans

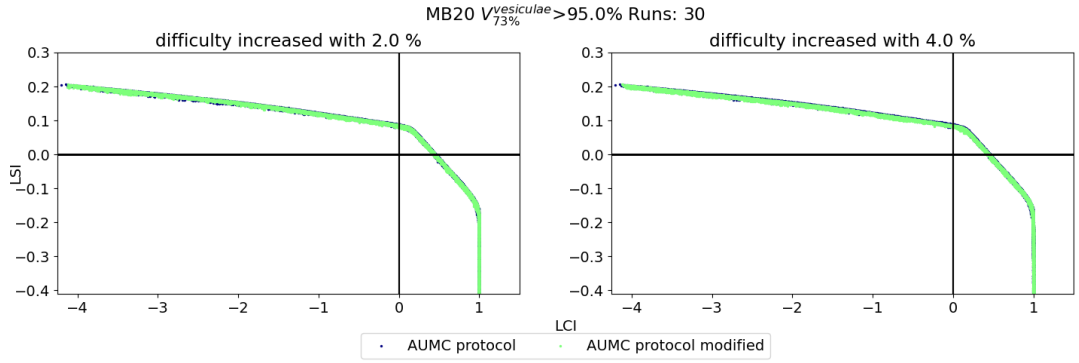

Slack analysis - Sparing indices - Patient MB21
LSI
LCI
$V^{ptv}_{150\%}$ 0 plans
$V^{ptv}_{200\%}$ 0 plans
$D^{bladder}_{1cm3}$ 0 plans
$D^{bladder}_{2cm3}$ 0 plans
D^{rectum}_{1cm3} 0 plans
D^{rectum}_{2cm3} 8 plans
$D^{urethra}_{30.0\%}$ 79 plans
$D^{urethra}_{0.1cm3}$ 0 plans

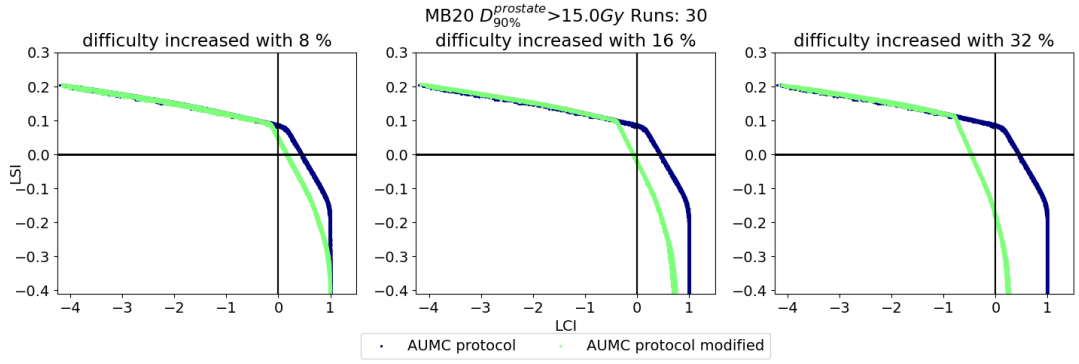

Slack analysis - Coverage indices - Patient MB22
LSI
LCI
$V^{ptv}_{100\%}$ 244 plans
$D^{ptv}_{90\%}$ 23 plans
$V^{vesiculae}_{73\%}$ 254 plans

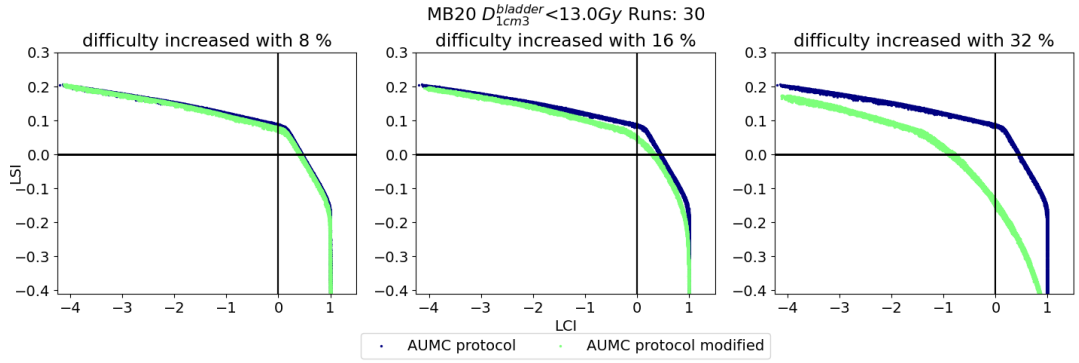

Slack analysis - Sparing indices - Patient MB22
LSI
LCI
$V^{ptv}_{150\%}$ 3 plans
$V^{ptv}_{200\%}$ 11 plans
$D^{bladder}_{1cm3}$ 180 plans
$D^{bladder}_{2cm3}$ 0 plans
D^{rectum}_{1cm3} 0 plans
D^{rectum}_{2cm3} 304 plans
$D^{urethra}_{30.0\%}$ 24 plans
$D^{urethra}_{0.1cm3}$ 0 plans

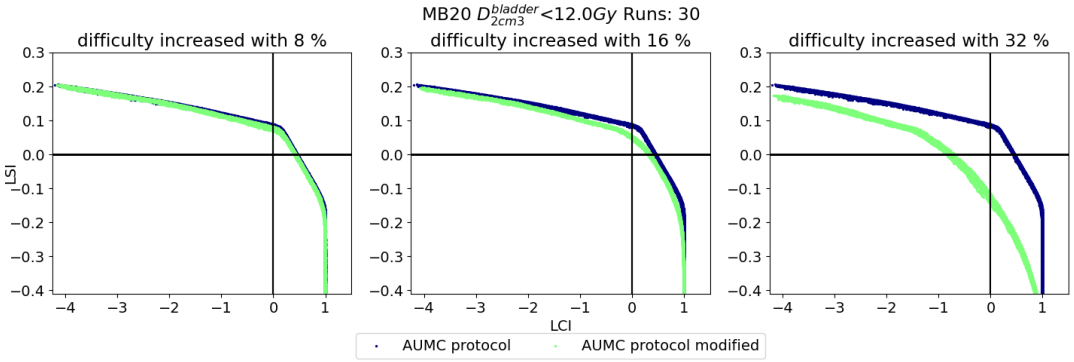

MB20 $V^{prostate}_{100\%}$>95.0% Runs: 30
difficulty increased with 2 %
difficulty increased with 4 %
LSI
LCI
AUMC protocol
AUMC protocol modified

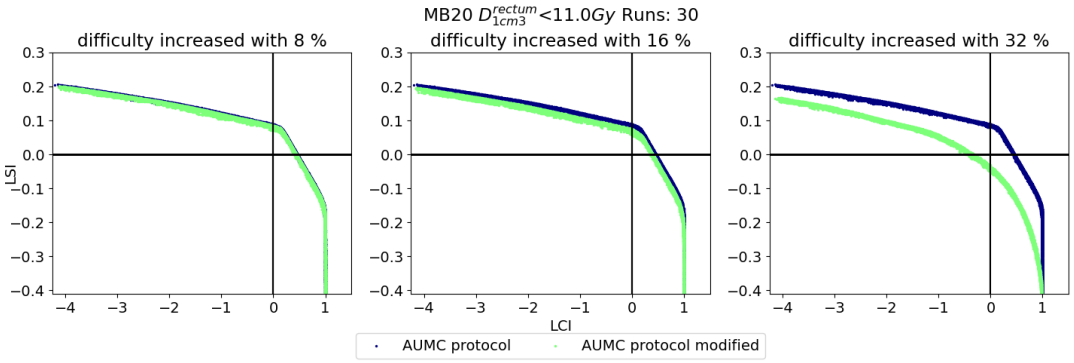

MB20 $V_{150\%}^{prostate}$<40.0% Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

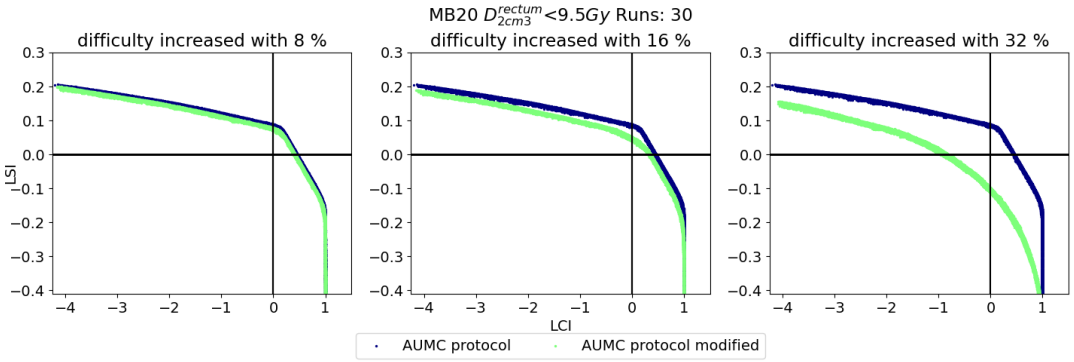

MB20 $V_{73\%}^{vesiculae}$>95.0% Runs: 30
difficulty increased with 2.0 %
difficulty increased with 4.0 %
LSI
LCI
AUMC protocol
AUMC protocol modified

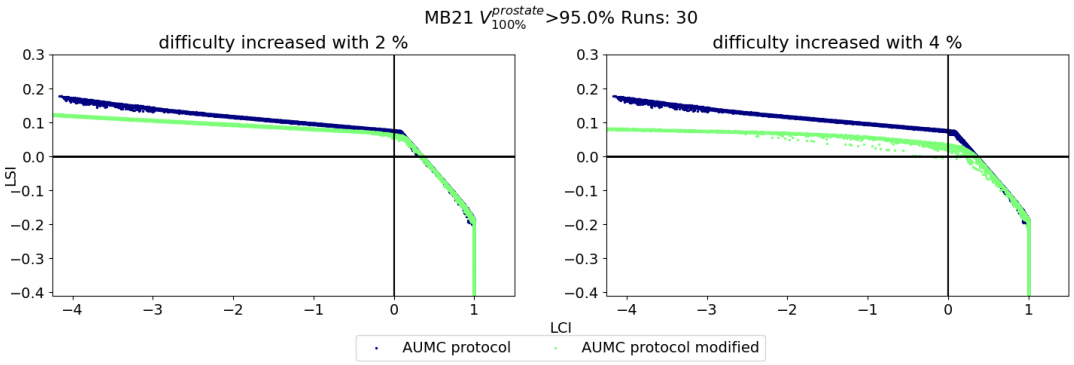

MB20 $D_{90\%}^{prostate}$>15.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

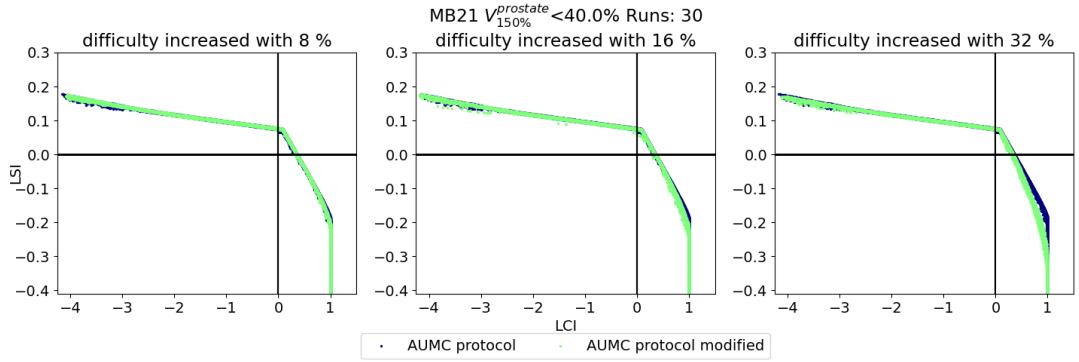

MB20 $D_{1cm3}^{bladder}$<13.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

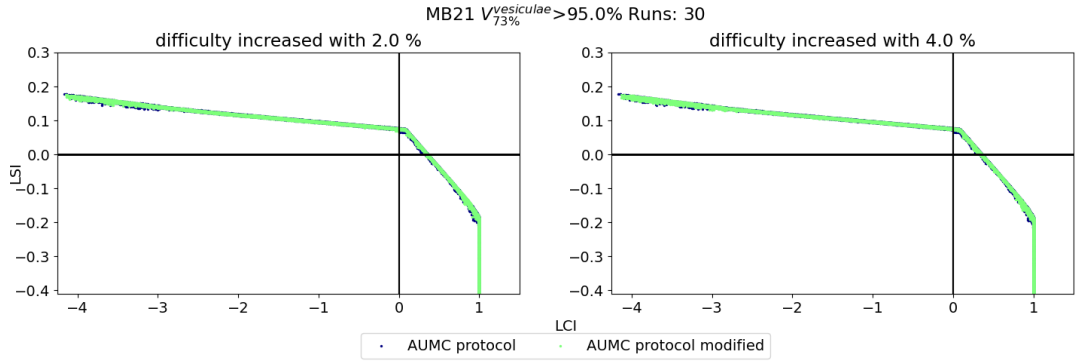

MB20 $D_{2cm3}^{bladder}$<12.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

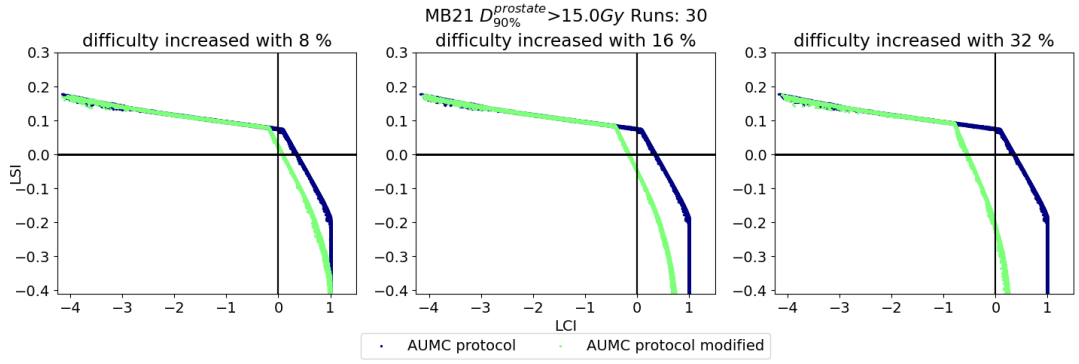

MB20 D_{1cm3}^{rectum}<11.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

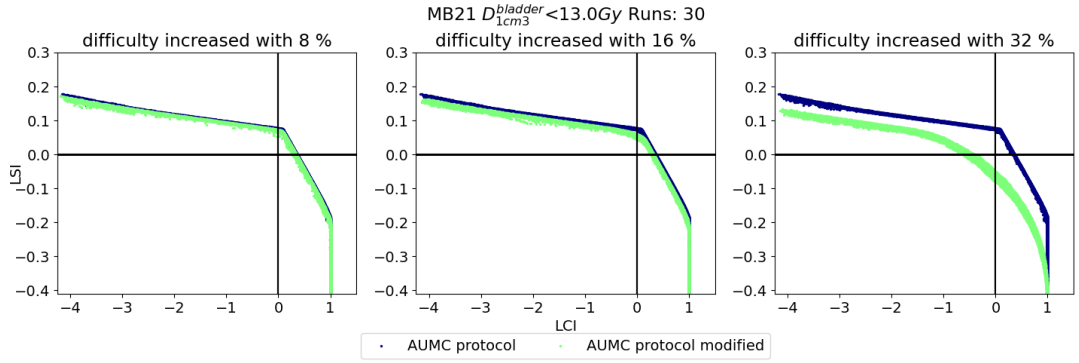

MB20 D_{2cm3}^{rectum}<9.5Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

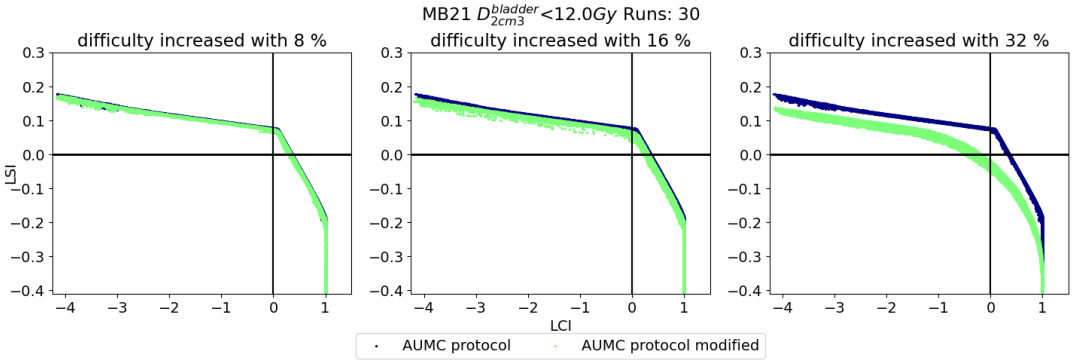

MB21 $V_{100\%}^{prostate}$>95.0% Runs: 30
difficulty increased with 2 %
difficulty increased with 4 %
LSI
LCI
AUMC protocol
AUMC protocol modified

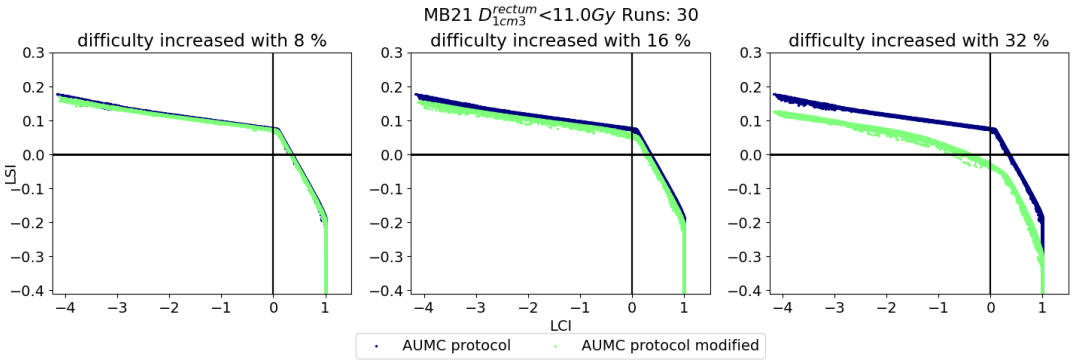

MB21 $V_{150\%}^{prostate}$<40.0% Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

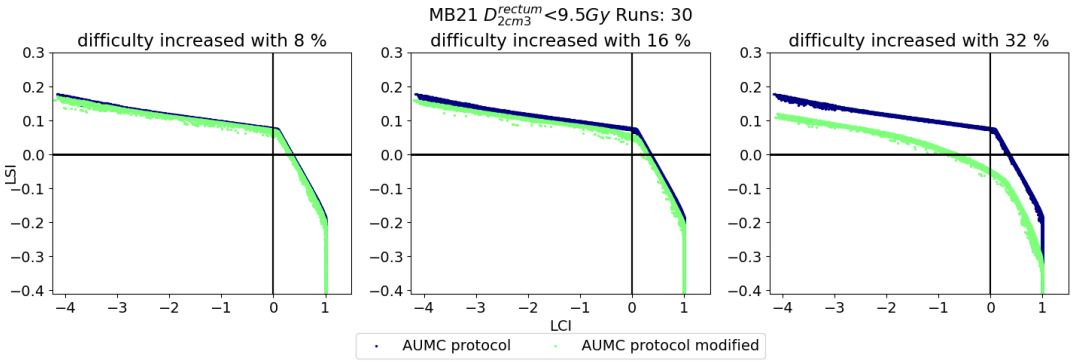

MB21 $V_{73\%}^{vesiculae}$>95.0% Runs: 30
difficulty increased with 2.0 %
difficulty increased with 4.0 %
LSI
LCI
AUMC protocol
AUMC protocol modified

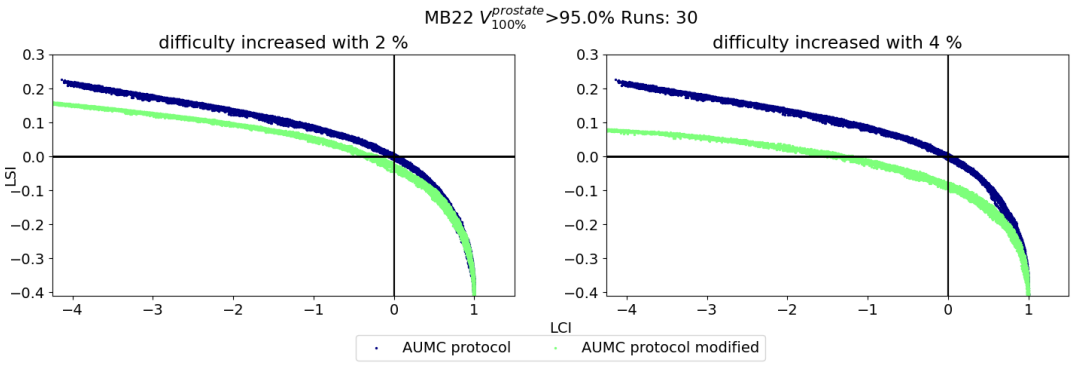

MB21 $D_{90\%}^{prostate}$>15.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

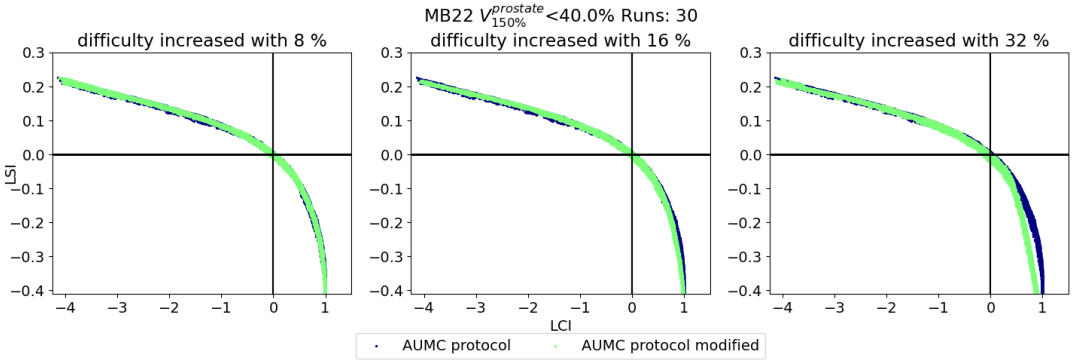

MB21 $D_{1cm3}^{bladder}$<13.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

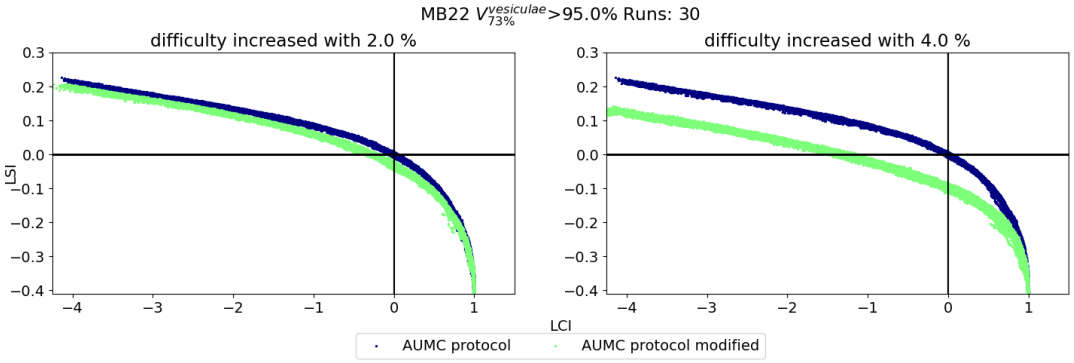

MB21 $D_{2cm3}^{bladder}$<12.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

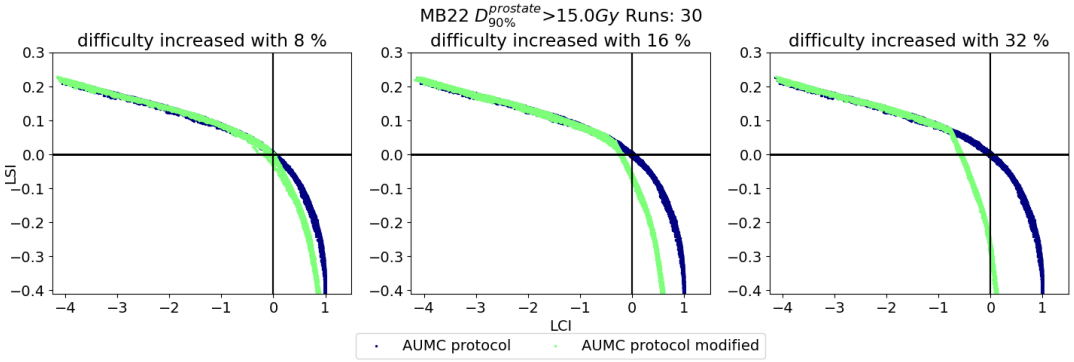

MB21 D_{1cm3}^{rectum}<11.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

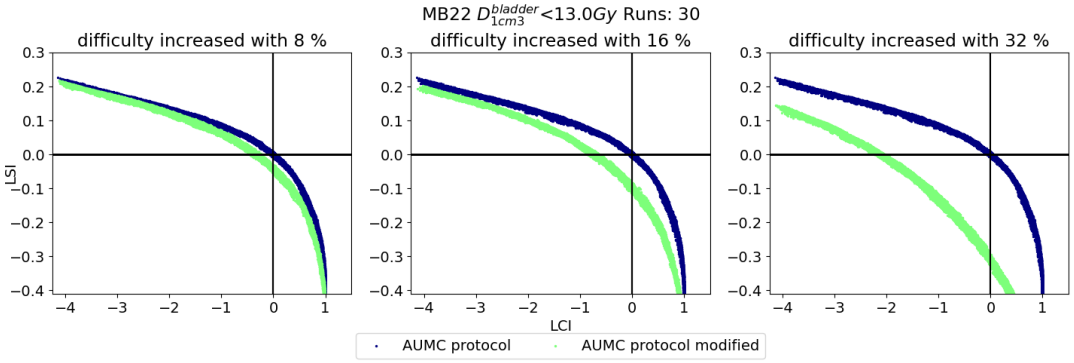

MB21 D_{2cm3}^{rectum}<9.5Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

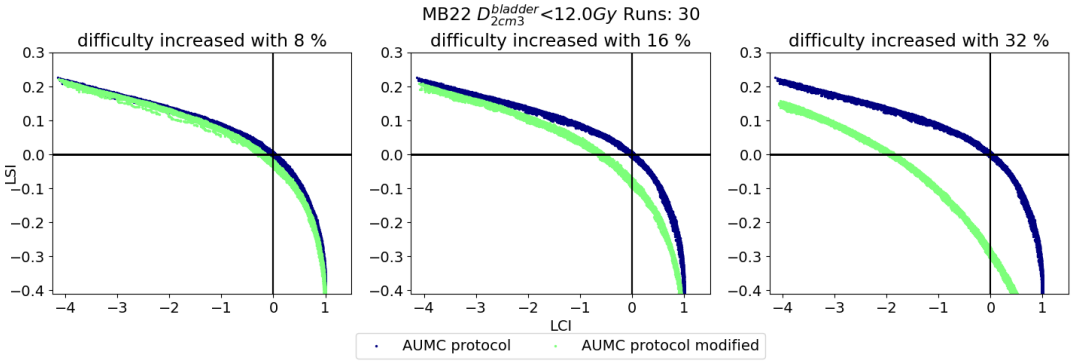

MB22 $V_{100\%}^{prostate}$>95.0% Runs: 30
difficulty increased with 2 %
difficulty increased with 4 %
LSI
LCI
AUMC protocol
AUMC protocol modified

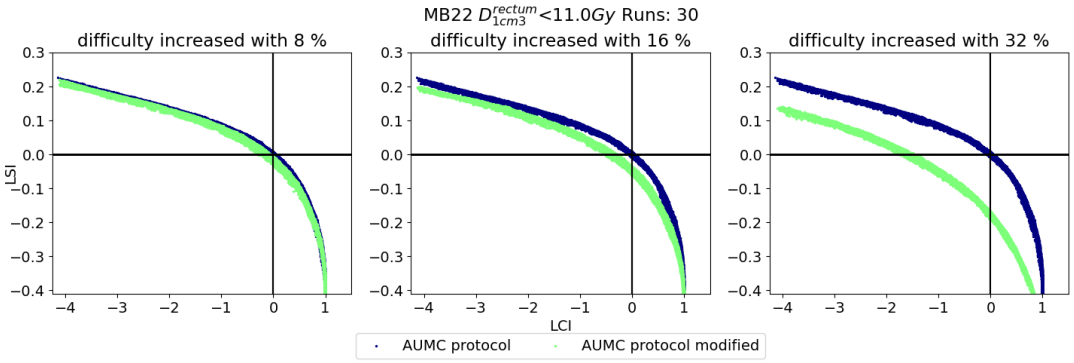

MB22 $V_{150\%}^{prostate}$<40.0% Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

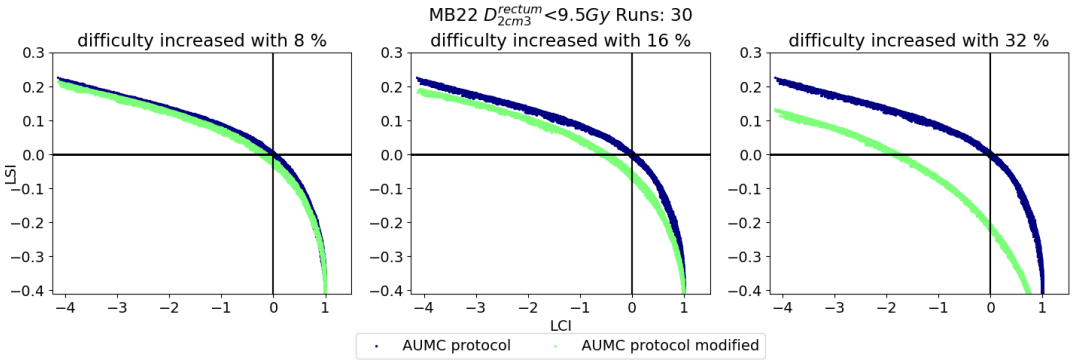

MB22 $V_{73\%}^{vesiculae}$>95.0% Runs: 30
difficulty increased with 2.0 %
difficulty increased with 4.0 %
LSI
LCI
AUMC protocol
AUMC protocol modified

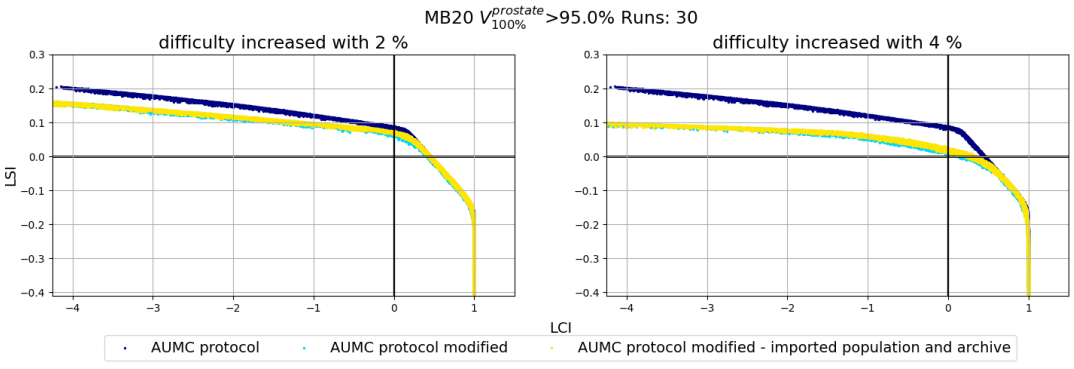

MB22 $D_{90\%}^{prostate}$>15.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

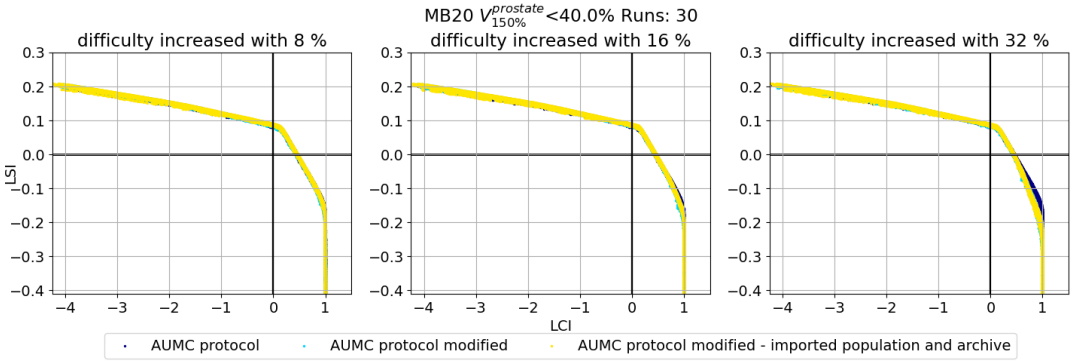

MB22 $D_{1cm3}^{bladder}$<13.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

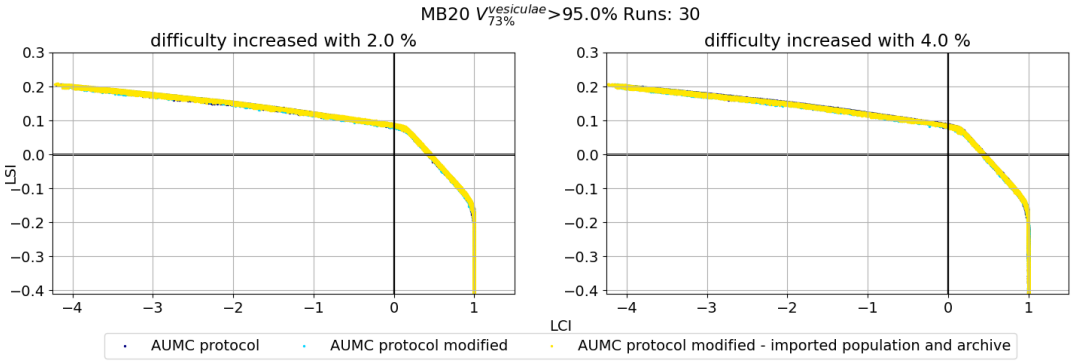

MB22 $D_{2cm3}^{bladder}$<12.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

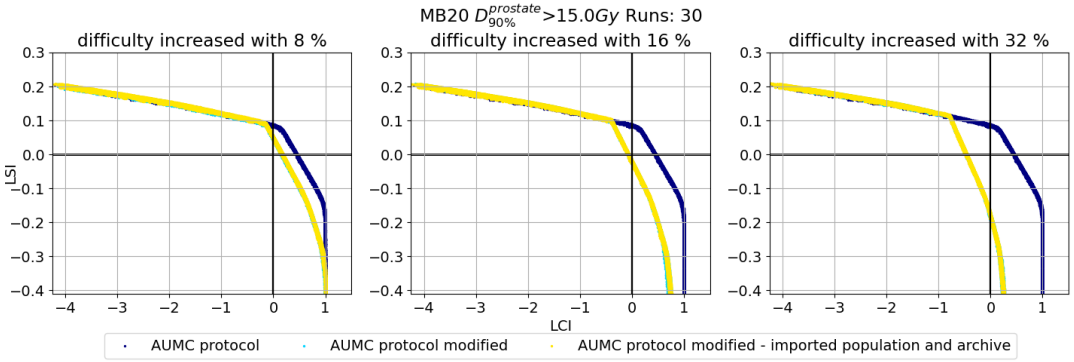

MB22 D_{1cm3}^{rectum}<11.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

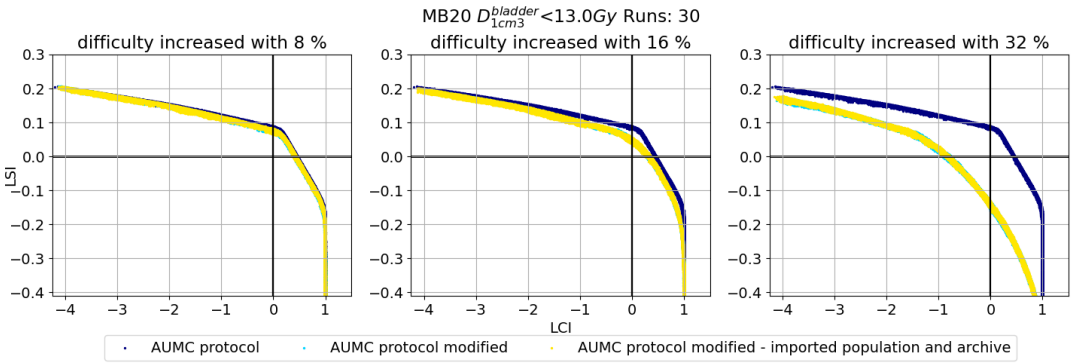

MB22 D_{2cm3}^{rectum}<9.5Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified

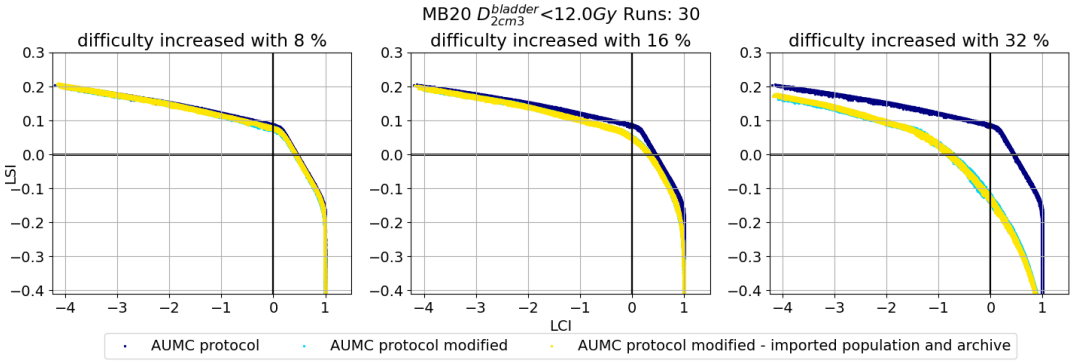

MB20 $V_{100\%}^{prostate}$>95.0% Runs: 30
difficulty increased with 2 %
difficulty increased with 4 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

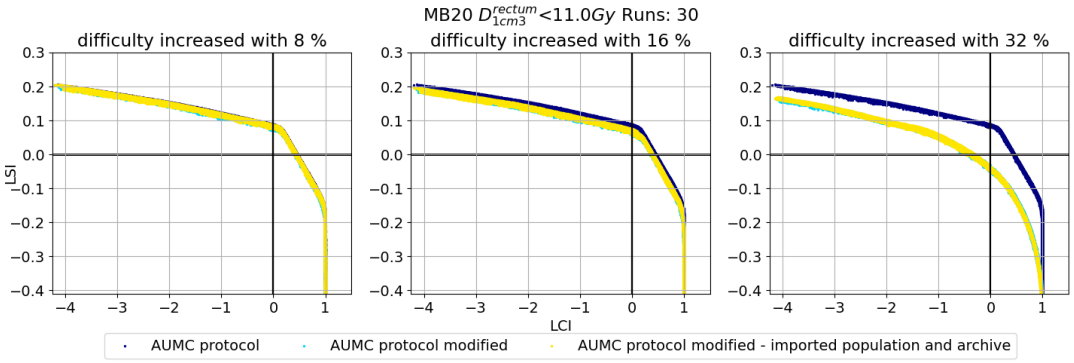

MB20 $V_{150\%}^{prostate}$<40.0% Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

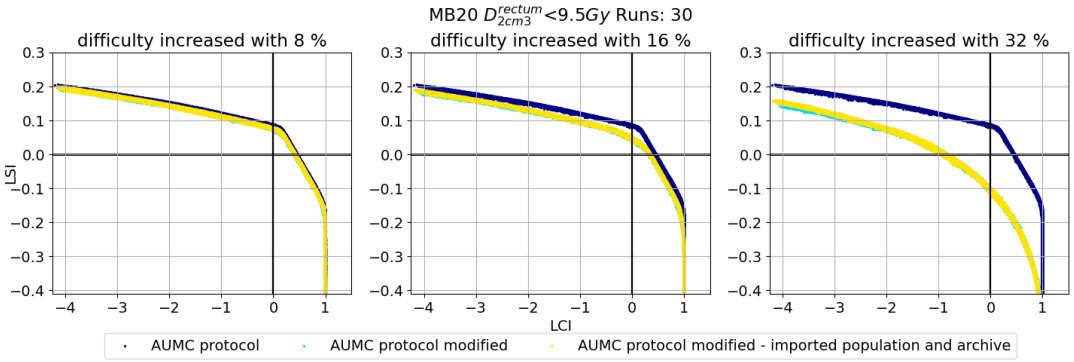

MB20 $V_{73\%}^{vesiculae}$>95.0% Runs: 30
difficulty increased with 2.0 %
difficulty increased with 4.0 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

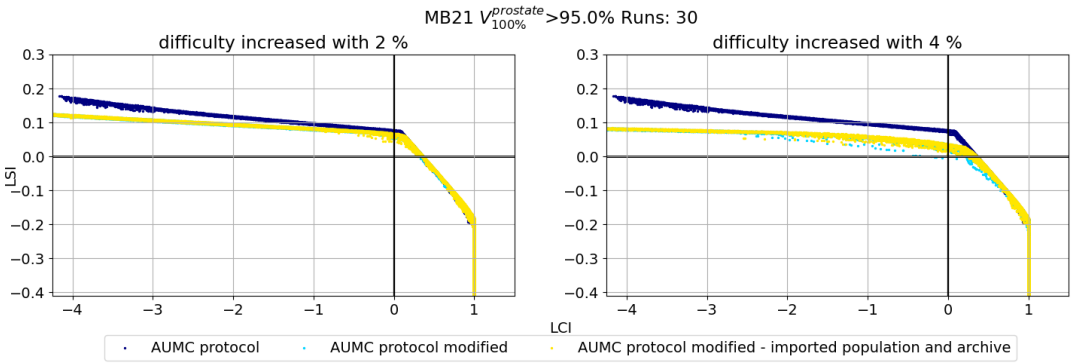

MB20 $D_{90\%}^{prostate}$>15.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

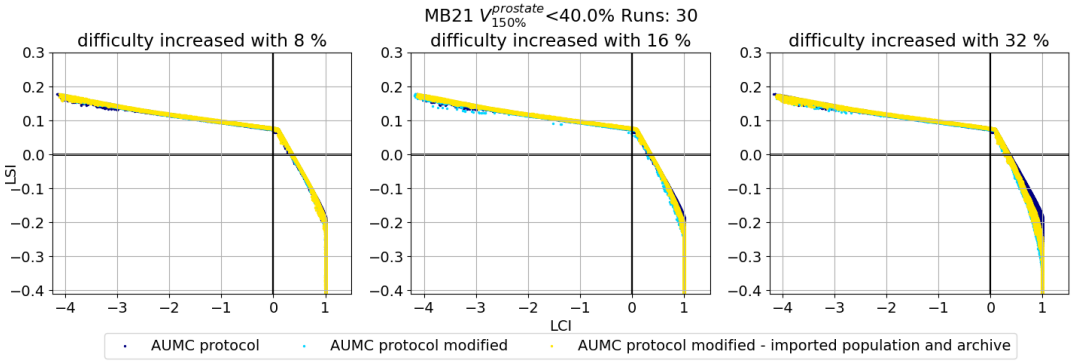

MB20 $D_{1cm3}^{bladder}$<13.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

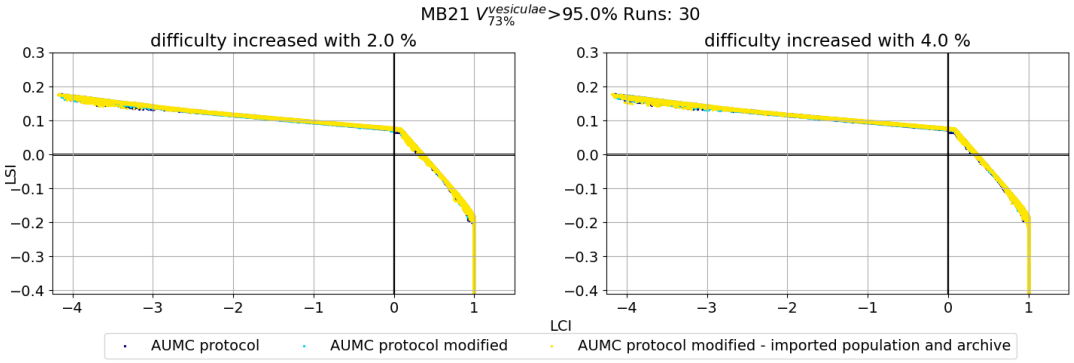

MB20 $D_{2cm3}^{bladder}$<12.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

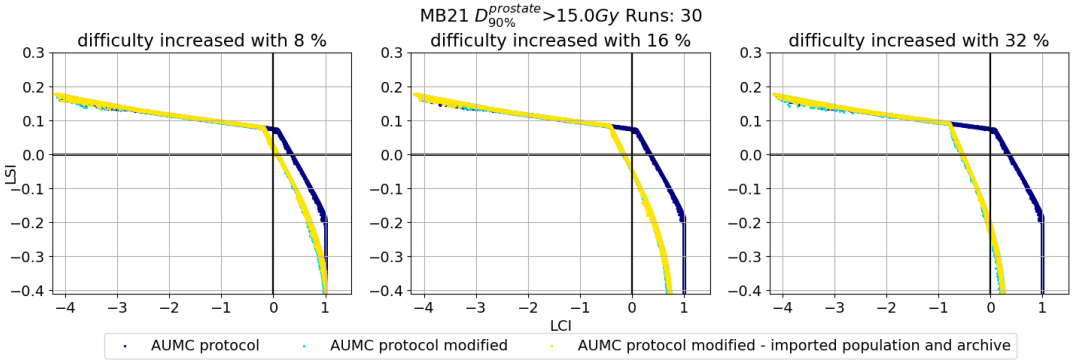

MB20 D_{1cm3}^{rectum}<11.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

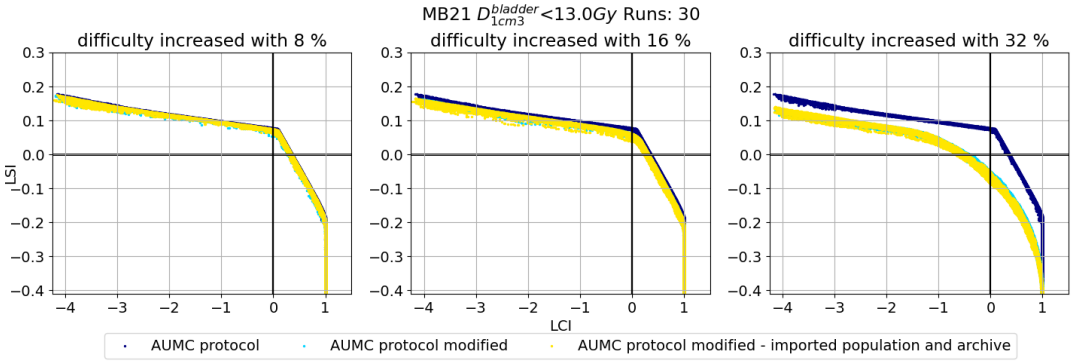

MB20 D_{2cm3}^{rectum}<9.5Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

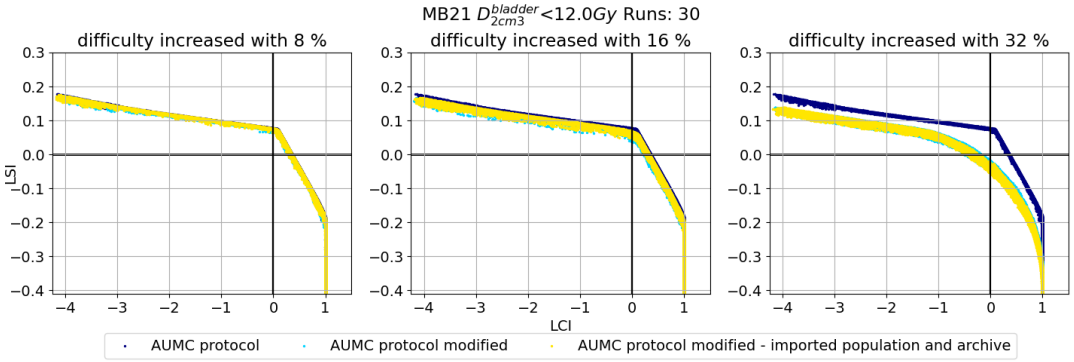

MB21 $V_{100\%}^{prostate}$>95.0% Runs: 30
difficulty increased with 2 %
difficulty increased with 4 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

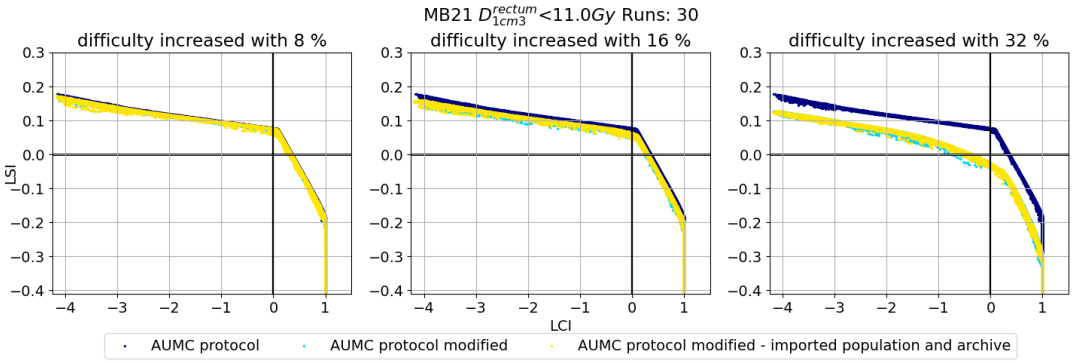

MB21 $V_{150\%}^{prostate}$<40.0% Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

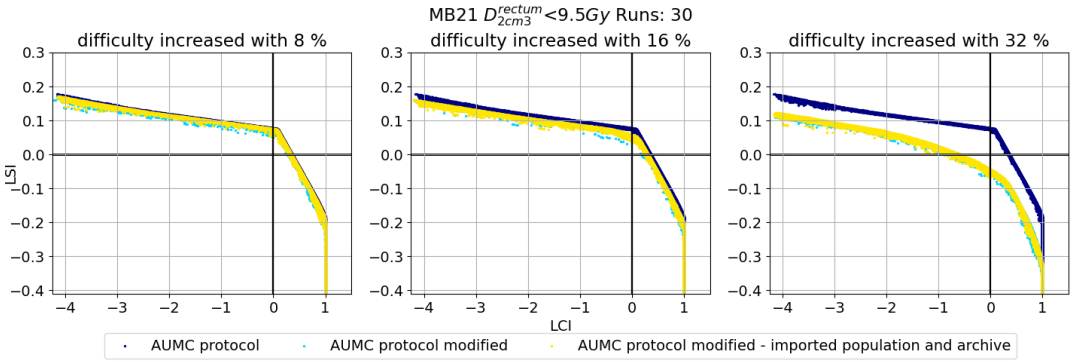

MB21 $V_{73\%}^{vesiculae}$>95.0% Runs: 30
difficulty increased with 2.0 %
difficulty increased with 4.0 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

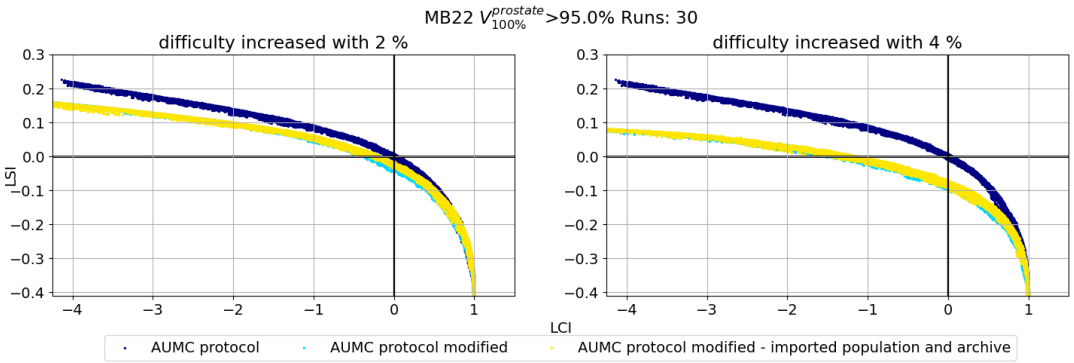

MB21 $D_{90\%}^{prostate}$>15.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

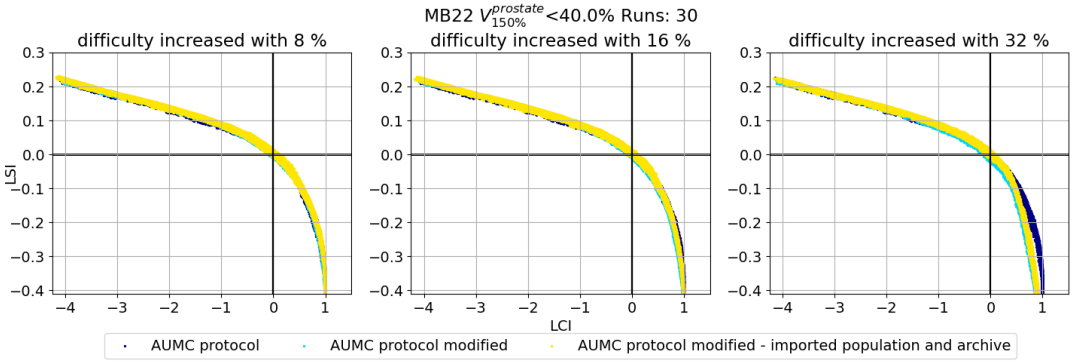

MB21 $D_{1cm3}^{bladder}$<13.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

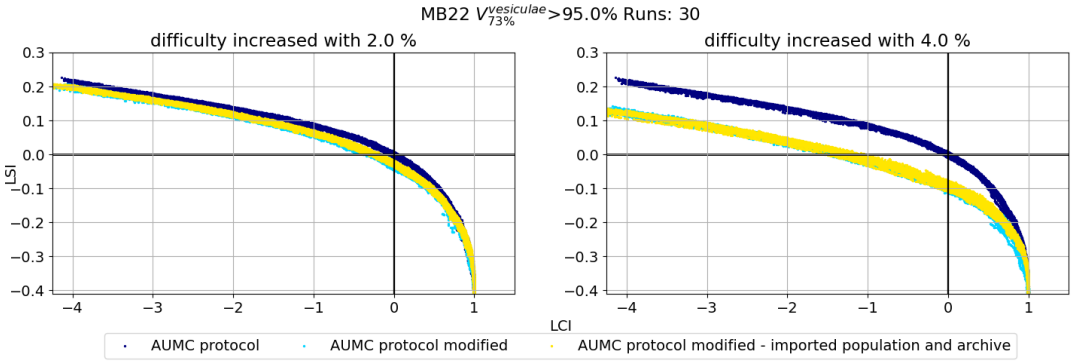

MB21 $D_{2cm3}^{bladder}$<12.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

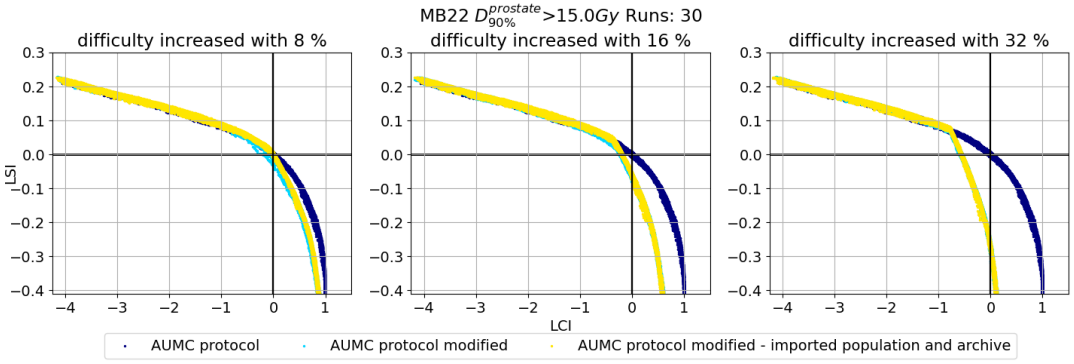

MB21 D_{1cm3}^{rectum}<11.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

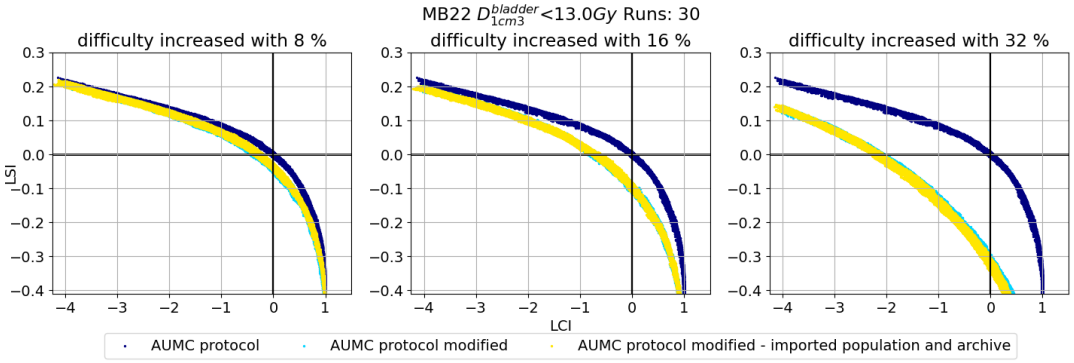

MB21 D_{2cm3}^{rectum}<9.5Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

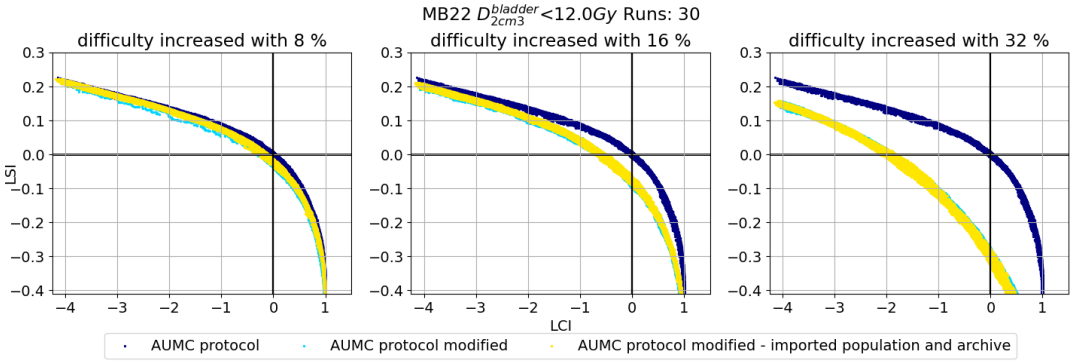

MB22 $V_{100\%}^{prostate}$>95.0% Runs: 30
difficulty increased with 2 %
difficulty increased with 4 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

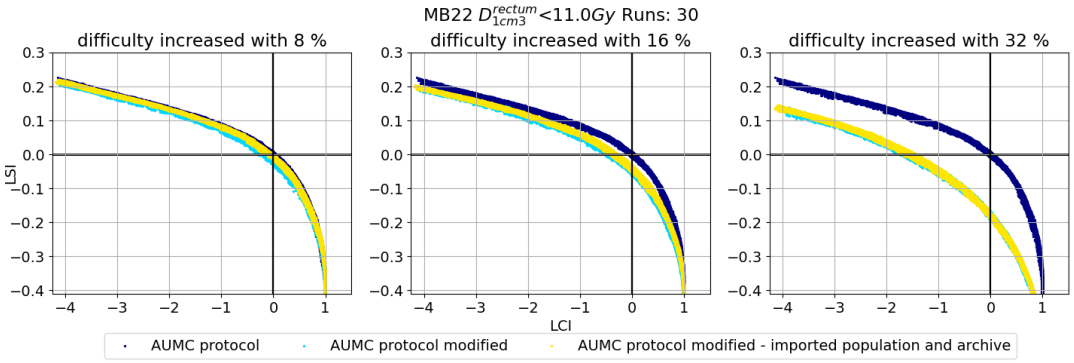

MB22 $V_{150\%}^{prostate}$<40.0% Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

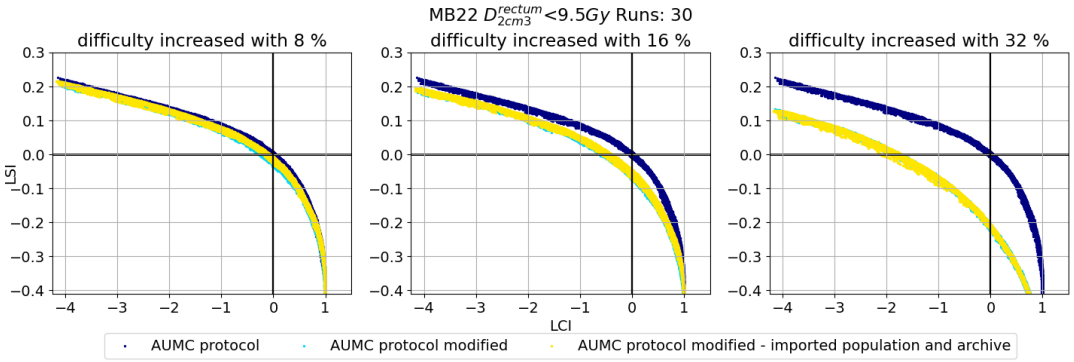

MB22 $V_{73\%}^{vesiculae}$>95.0% Runs: 30
difficulty increased with 2.0 %
difficulty increased with 4.0 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

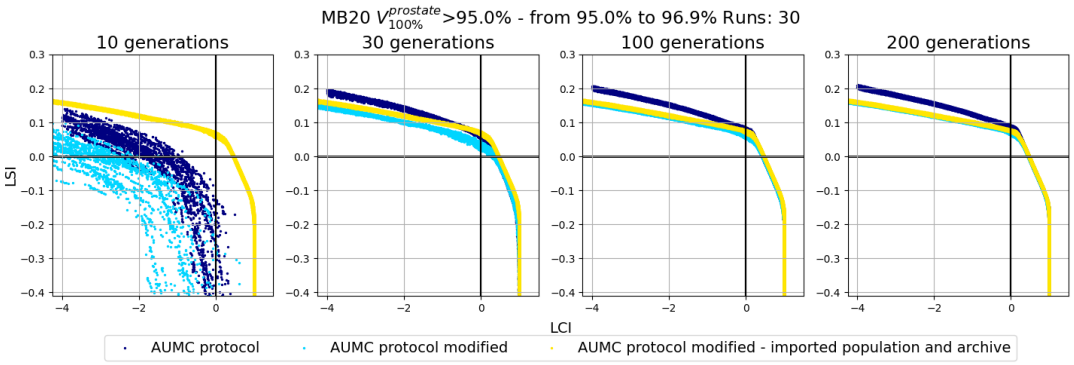

MB22 $D_{90\%}^{prostate}$>15.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

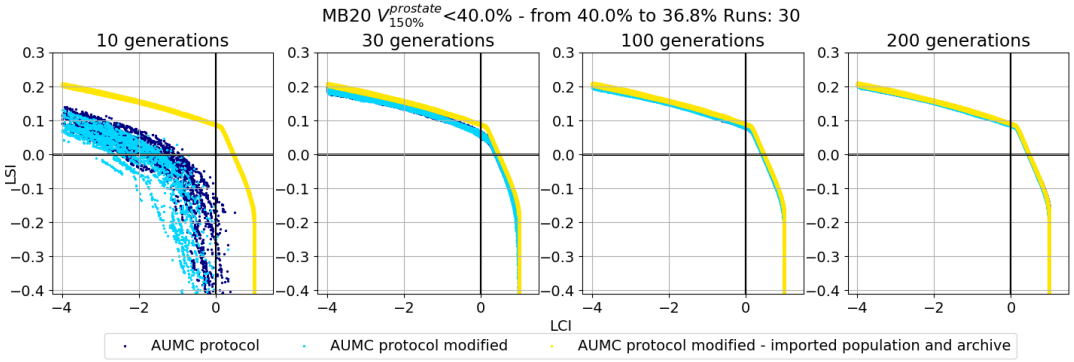

MB22 $D_{1cm3}^{bladder}$<13.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

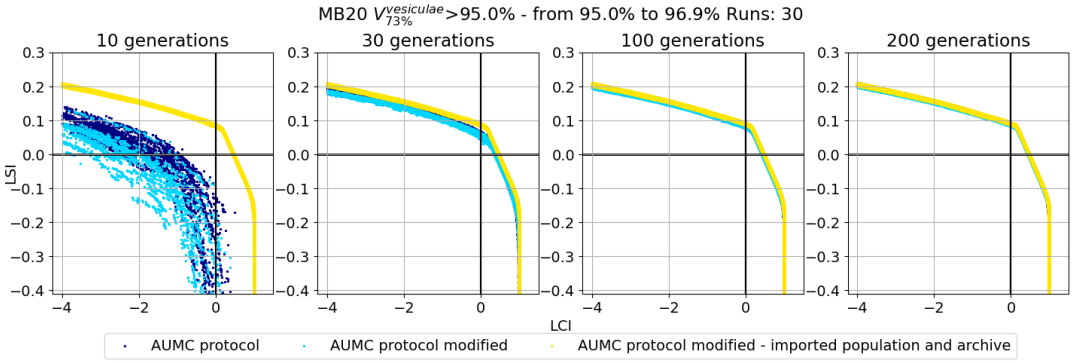

MB22 $D_{2cm3}^{bladder}$<12.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

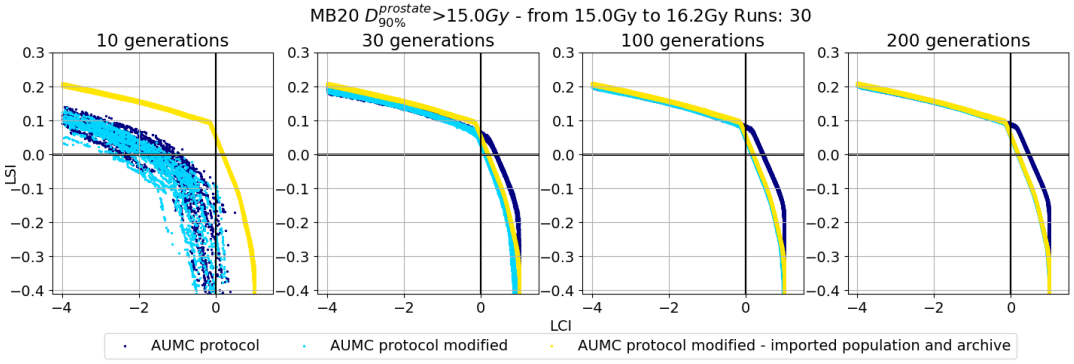

MB22 D_{1cm3}^{rectum}<11.0Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

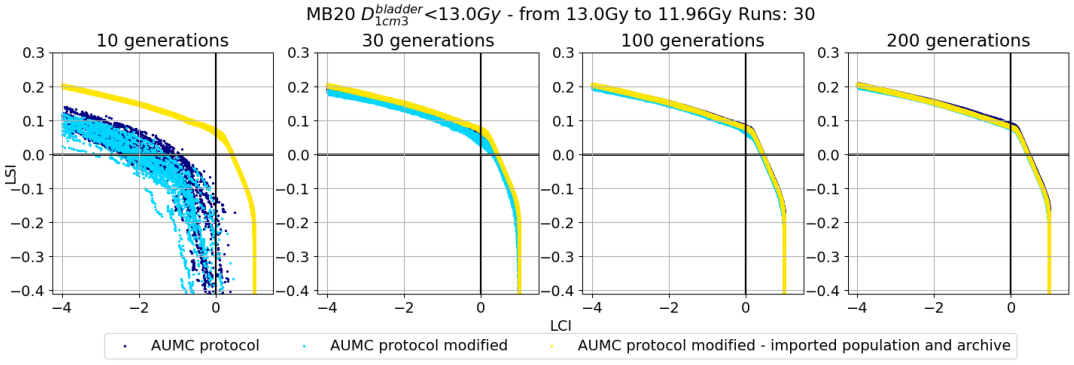

MB22 D_{2cm3}^{rectum}<9.5Gy Runs: 30
difficulty increased with 8 %
difficulty increased with 16 %
difficulty increased with 32 %
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

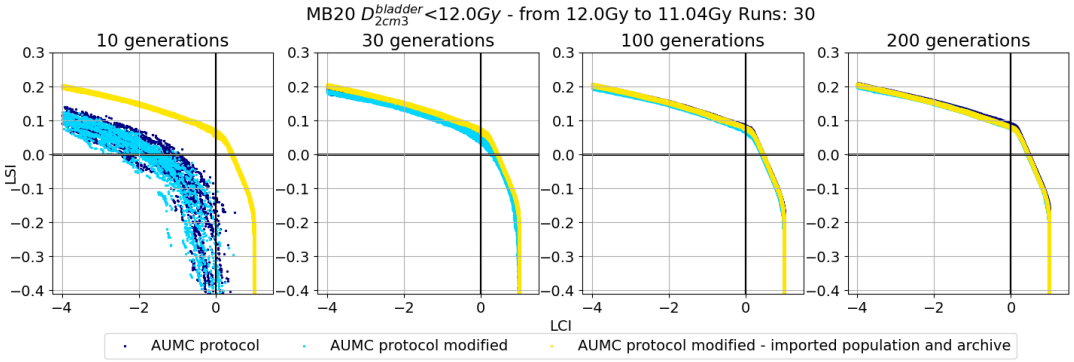

MB20 $V_{100\%}^{prostate}$>95.0% - from 95.0% to 96.9% Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

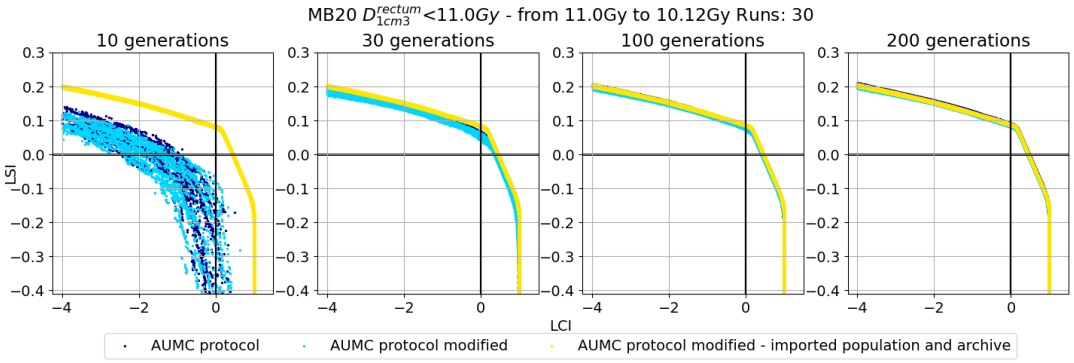

MB20 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 36.8% Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

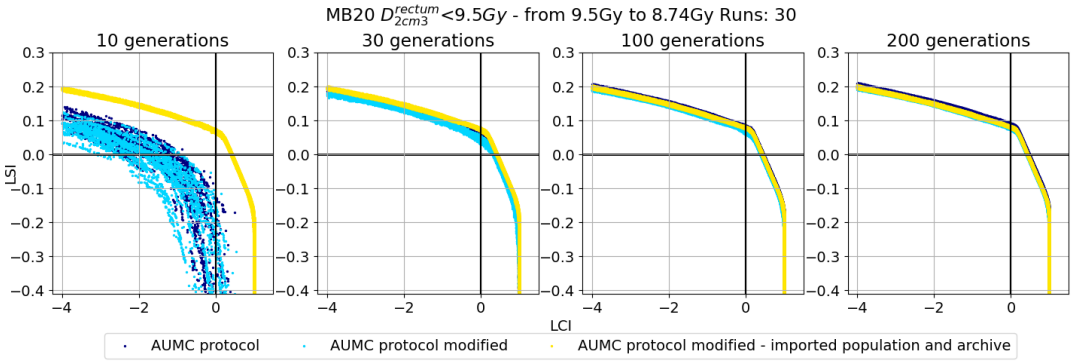

MB20 $V_{73\%}^{vesiculae}$>95.0% - from 95.0% to 96.9% Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

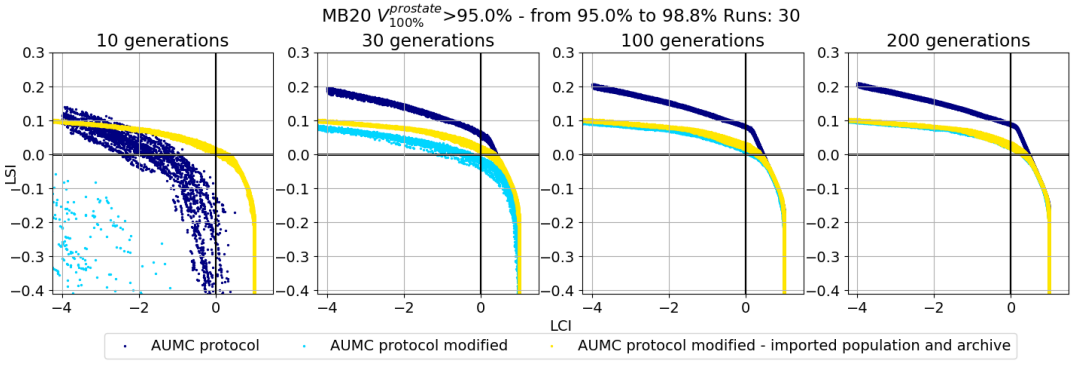

MB20 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 16.2Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

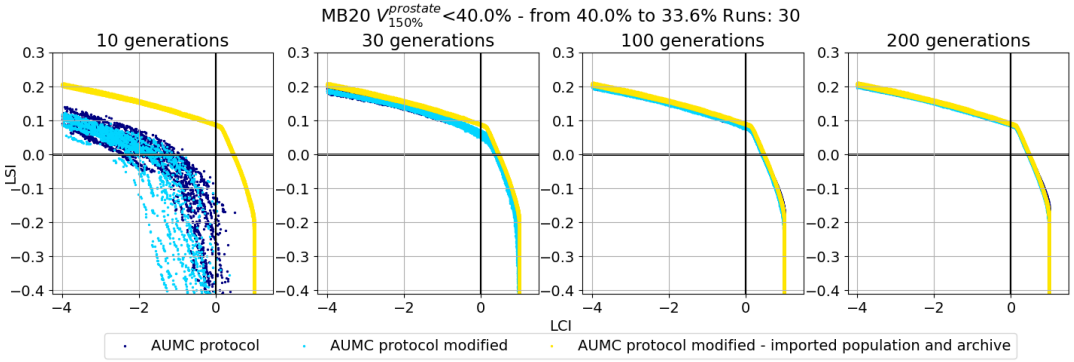

MB20 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 11.96Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

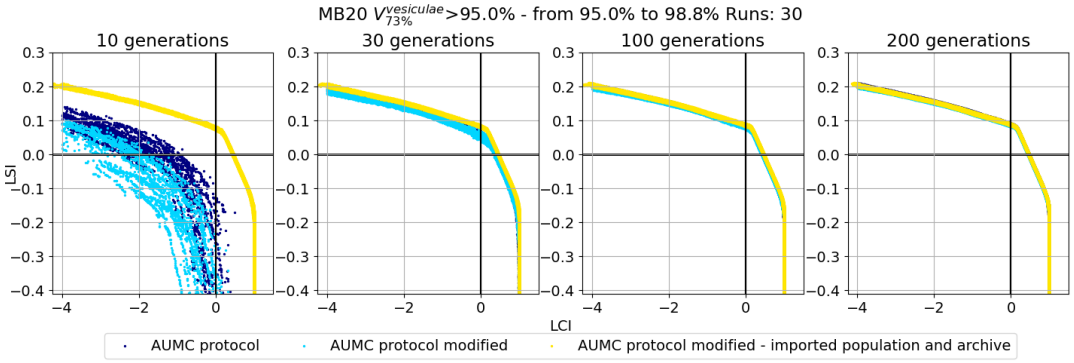

MB20 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 11.04Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

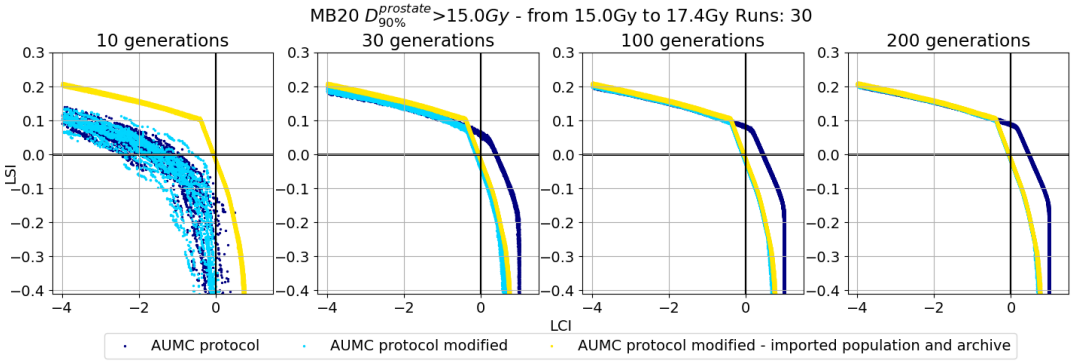

MB20 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 10.12Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

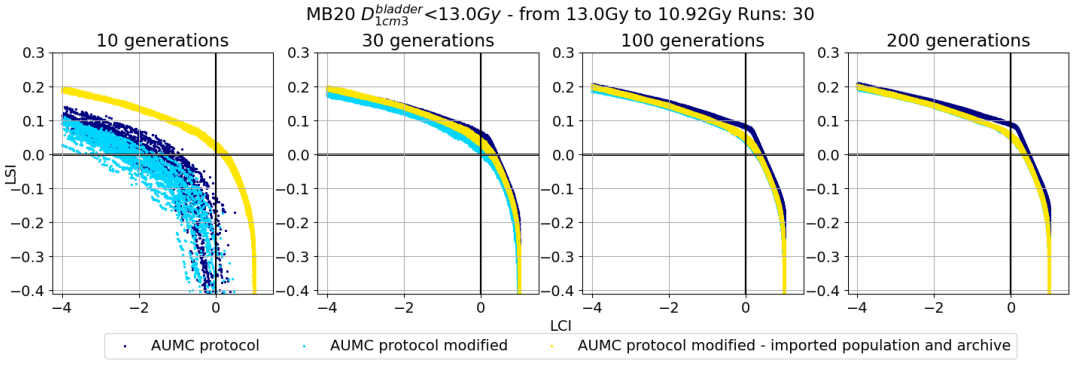

MB20 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 8.74Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

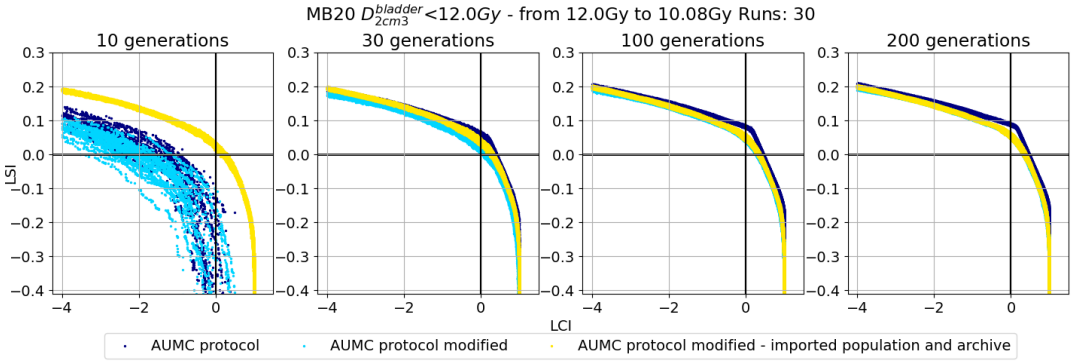

MB20 $V_{100\%}^{prostate}$>95.0% - from 95.0% to 98.8% Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

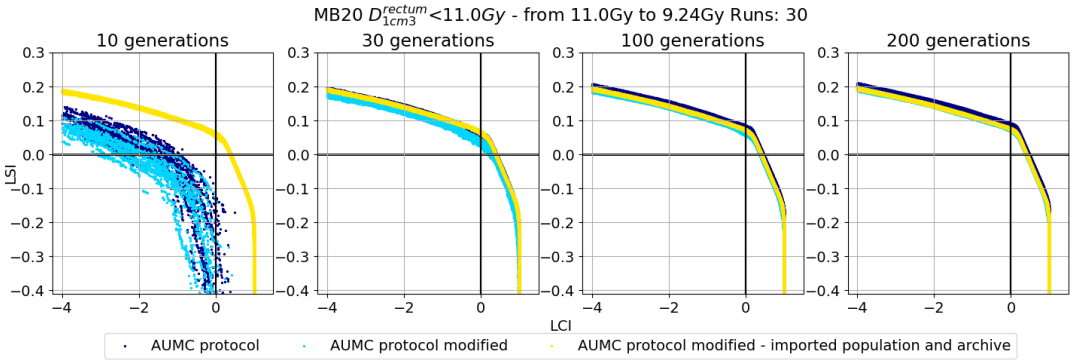

MB20 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 33.6% Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

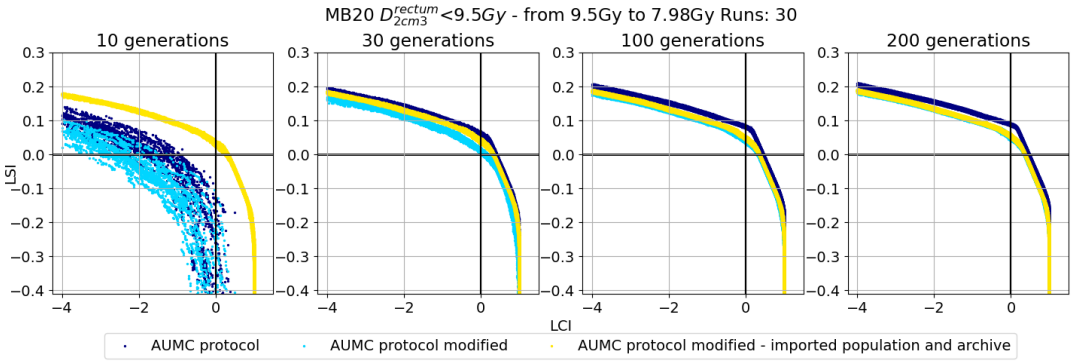

MB20 $V_{73\%}^{vesiculae}$>95.0% - from 95.0% to 98.8% Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

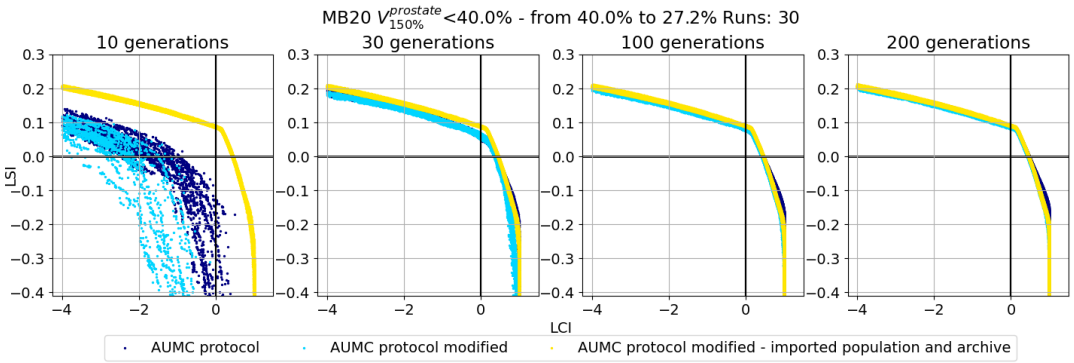

MB20 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 17.4Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

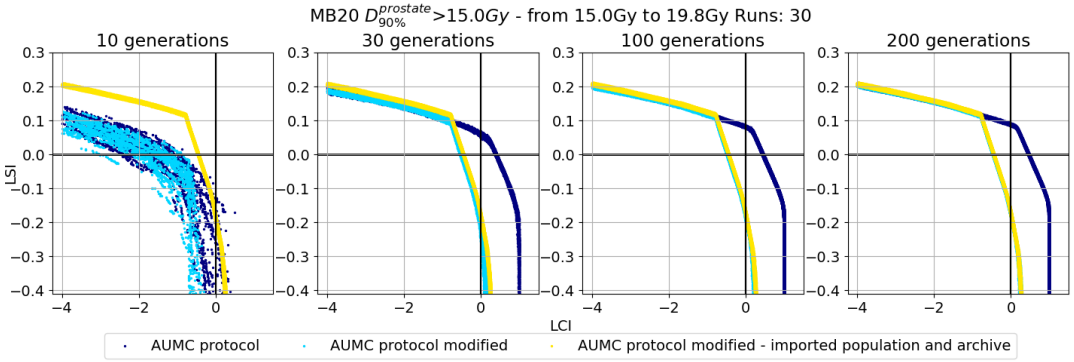

MB20 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 10.92Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

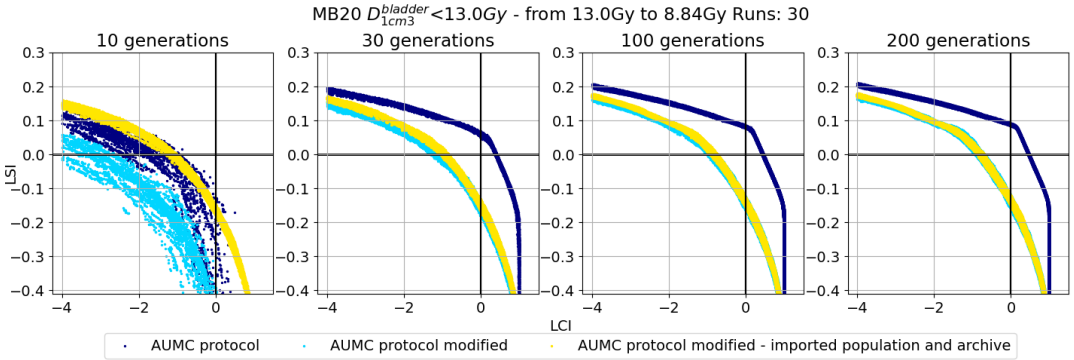

MB20 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 10.08Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

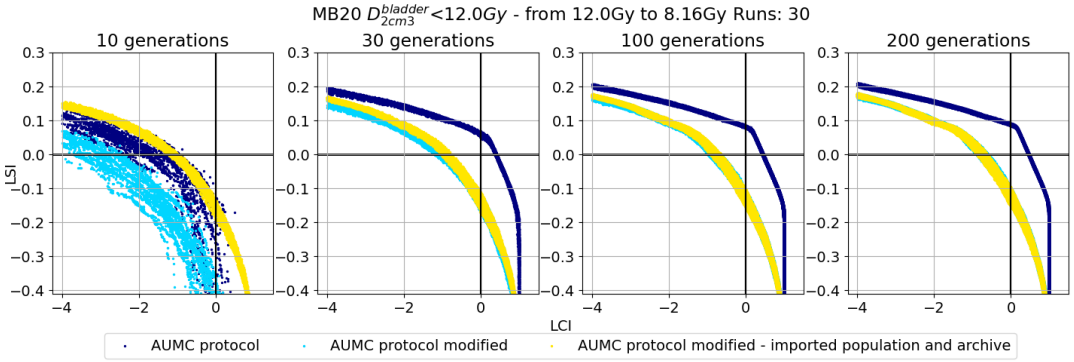

MB20 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 9.24Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

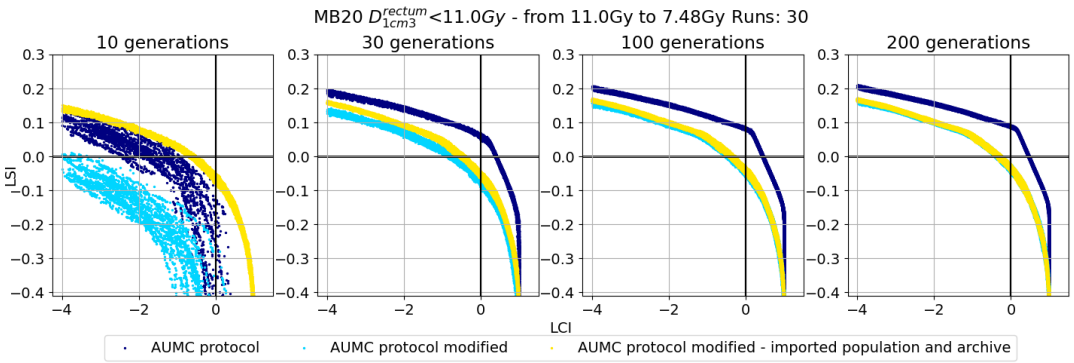

MB20 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 7.98Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

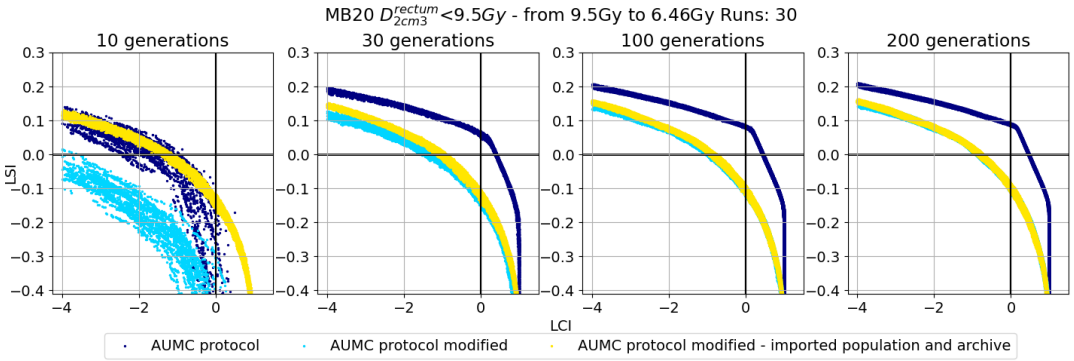

MB20 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 27.2% Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

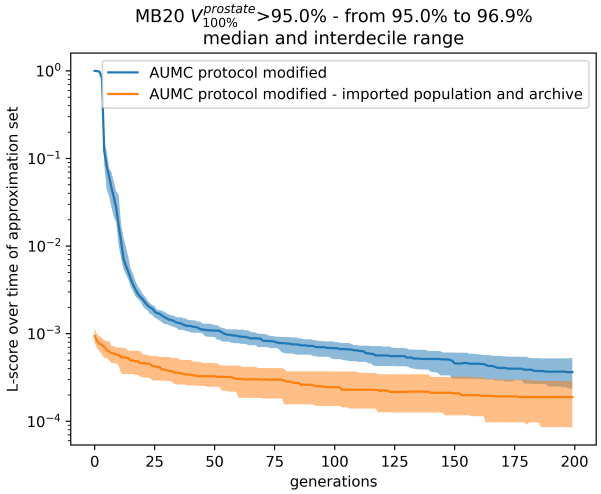

MB20 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 19.8Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

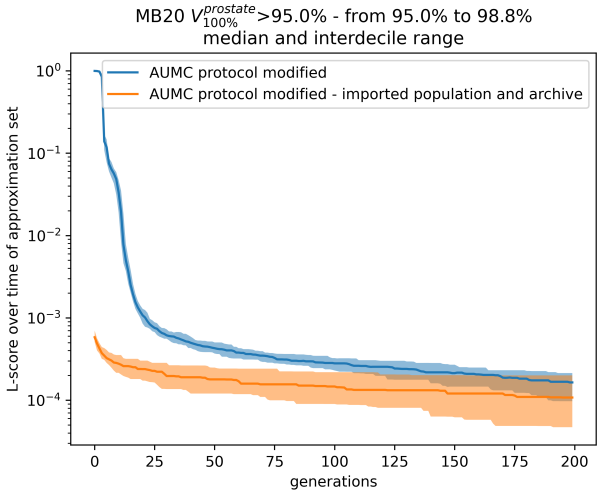

MB20 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 8.84Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

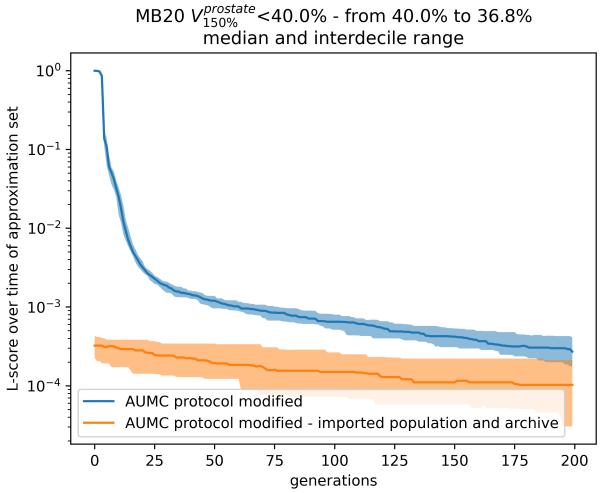

MB20 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 8.16Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

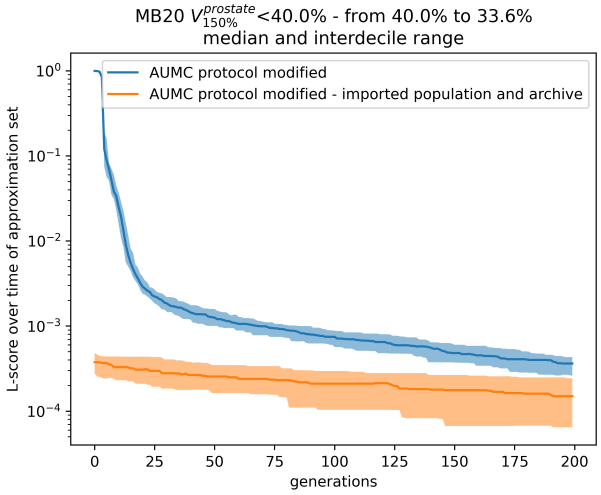

MB20 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 7.48Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

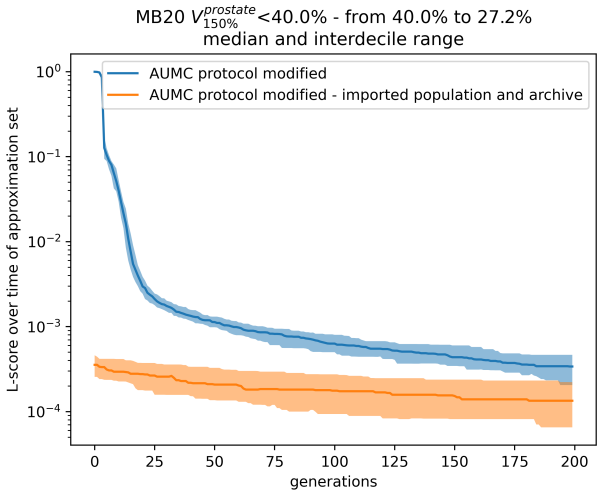

MB20 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 6.46Gy Runs: 30
10 generations
30 generations
100 generations
200 generations
LSI
LCI
AUMC protocol
AUMC protocol modified
AUMC protocol modified - imported population and archive

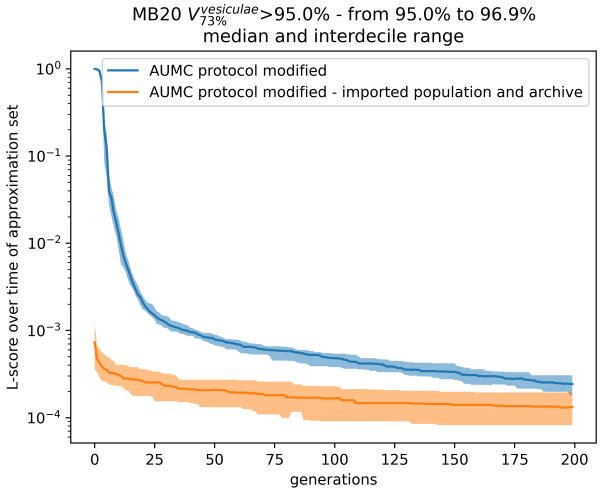

MB20 $V_{100\%}^{prostate}$>95.0% - from 95.0% to 96.9%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

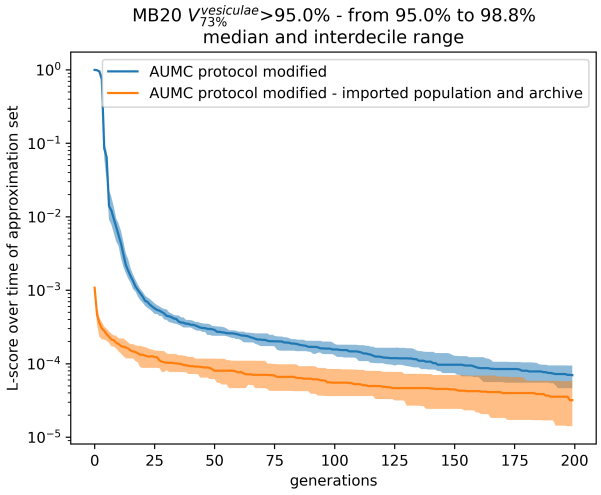

MB20 $V_{100\%}^{prostate}$>95.0% - from 95.0% to 98.8%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

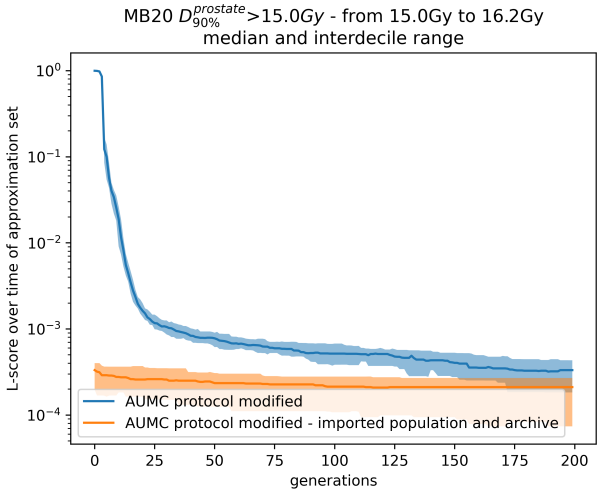

MB20 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 36.8%
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

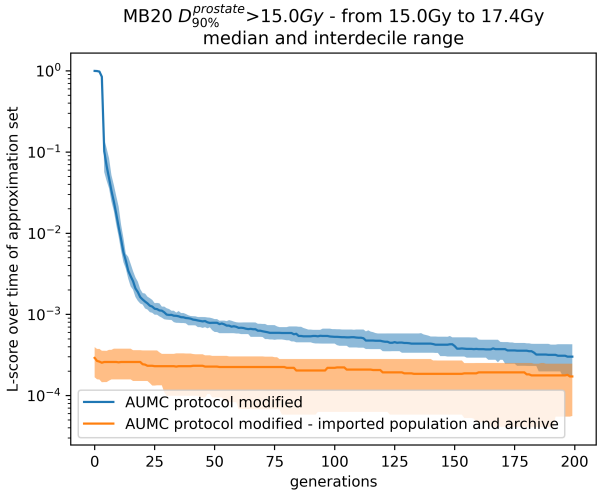

MB20 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 33.6%
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

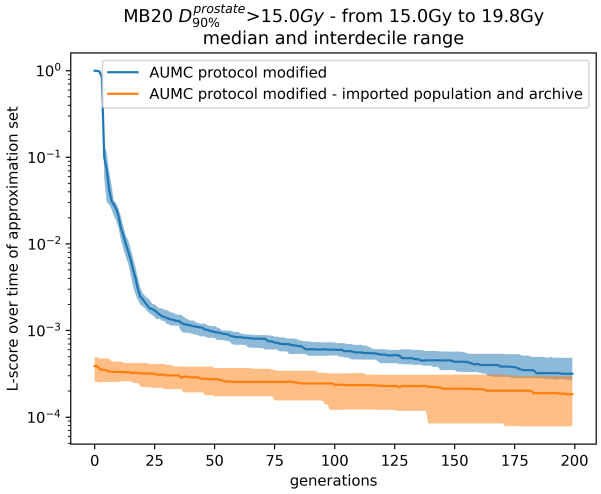

MB20 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 27.2%
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

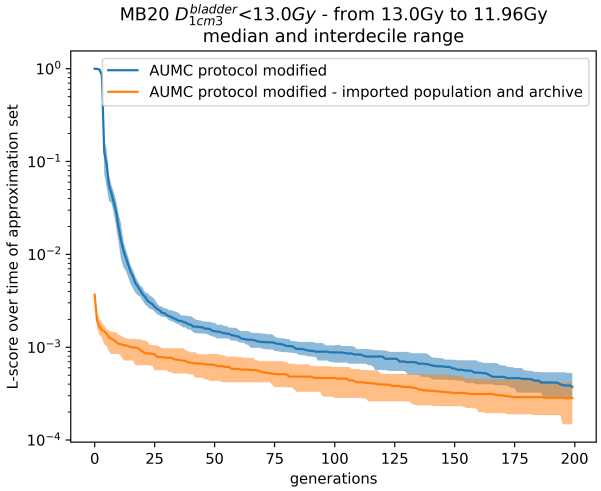

MB20 $V^{vesiculae}_{73\%}$>95.0% - from 95.0% to 96.9%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

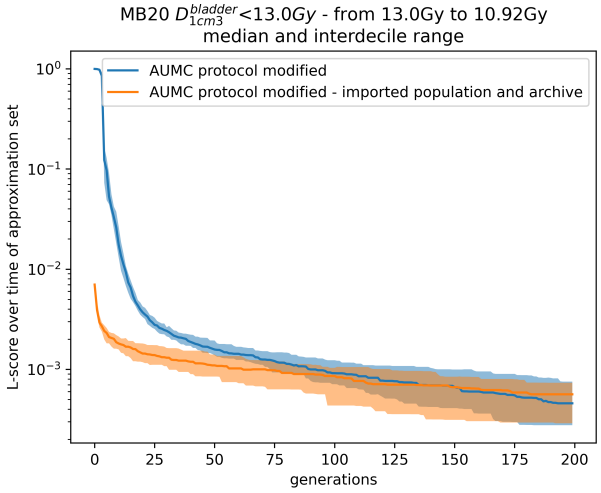

MB20 $V^{vesiculae}_{73\%}$>95.0% - from 95.0% to 98.8%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

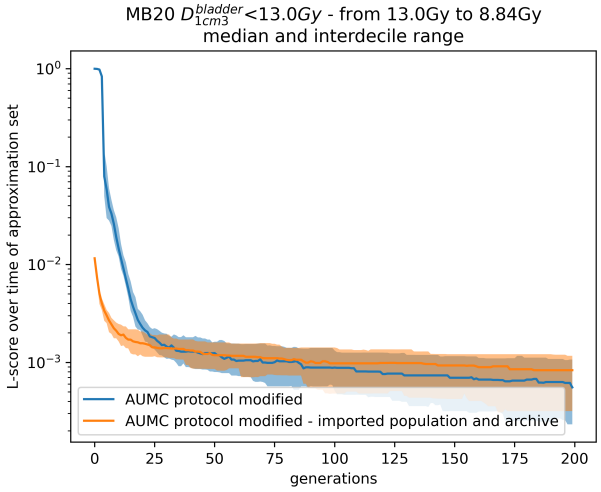

MB20 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 16.2Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

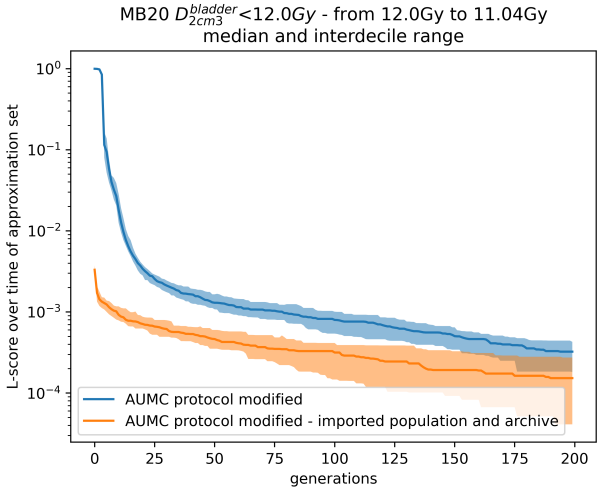

MB20 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 17.4Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

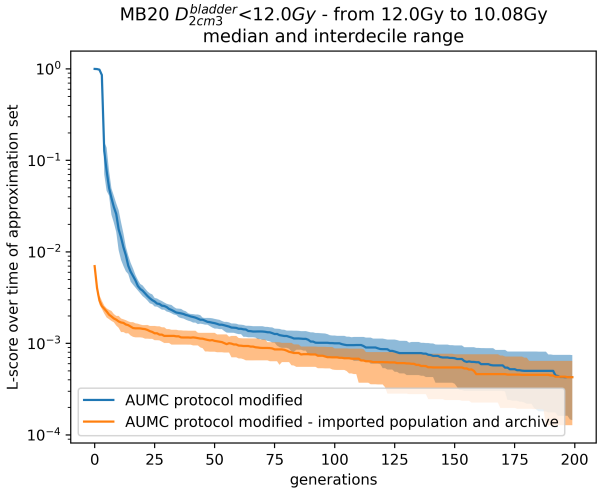

MB20 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 19.8Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

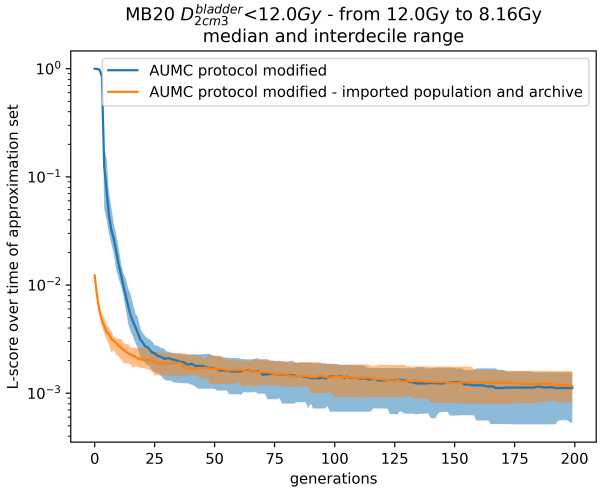

MB20 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 11.96Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

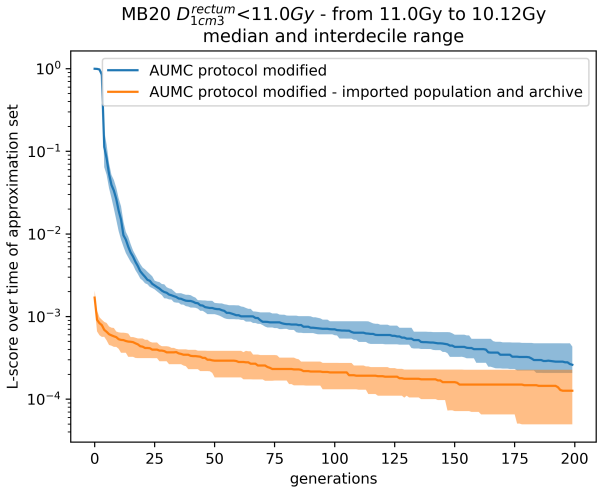

MB20 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 10.92Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

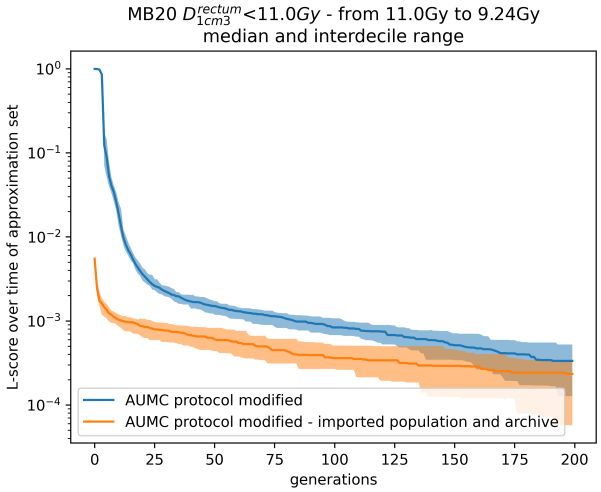

MB20 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 8.84Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

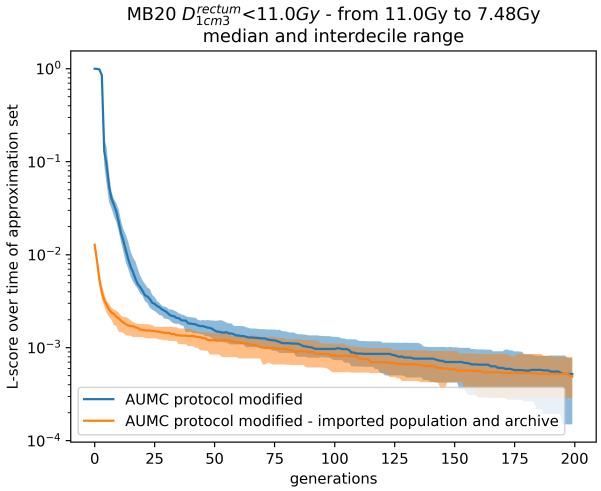

MB20 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 11.04Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

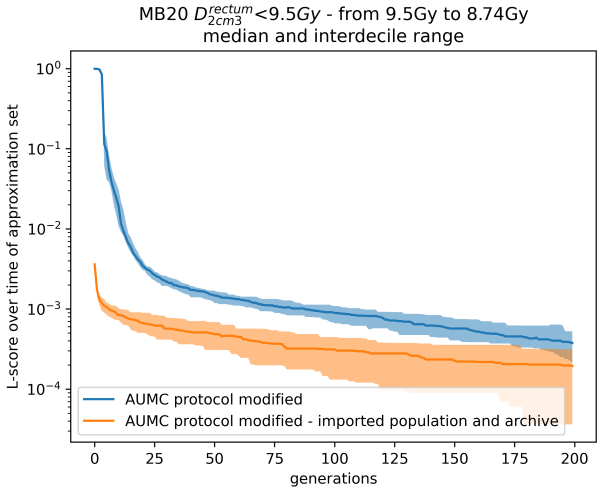

MB20 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 10.08Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

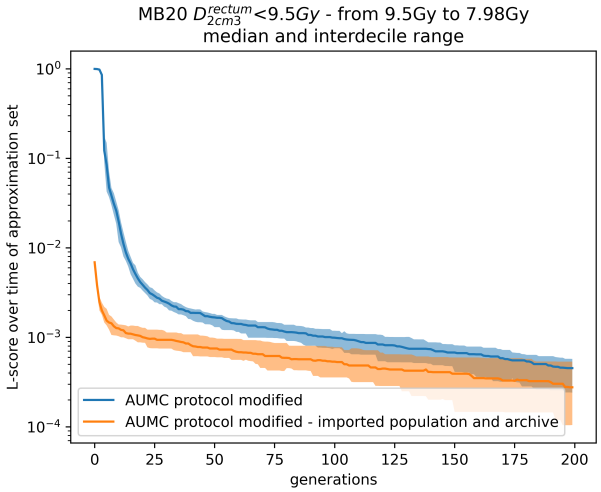

MB20 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 8.16Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

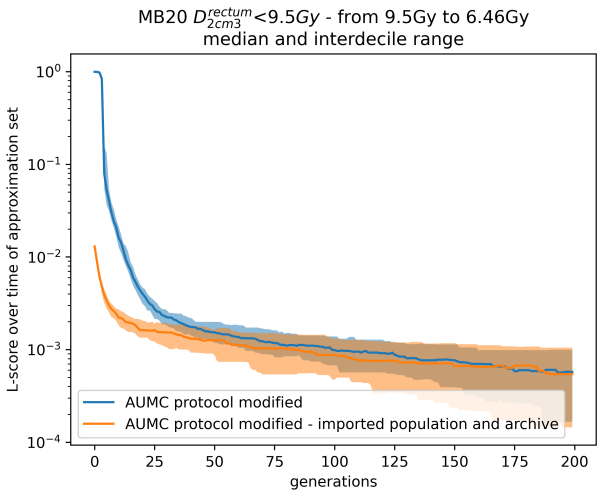

MB20 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 10.12Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

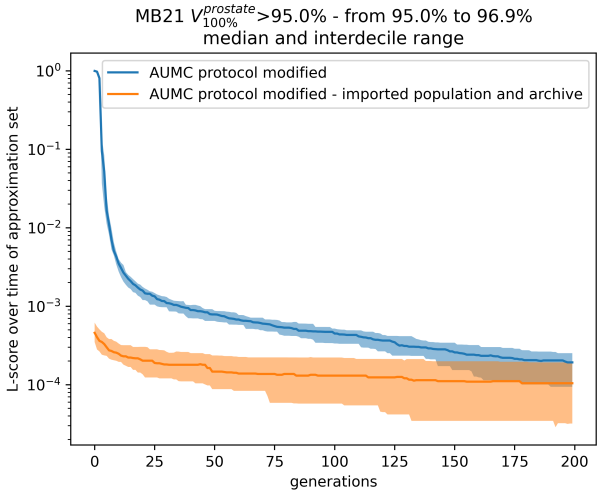

MB20 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 9.24Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

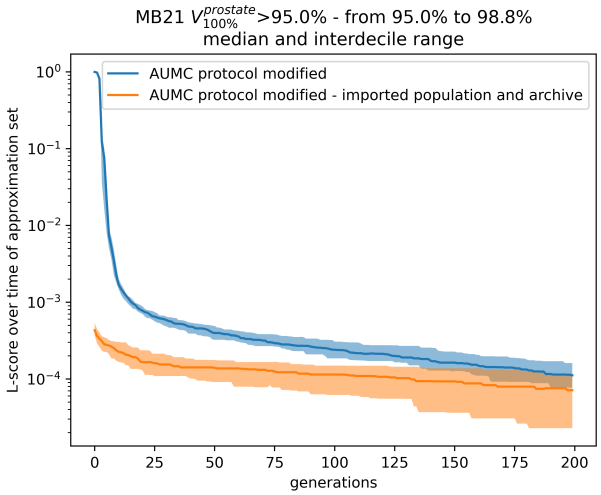

MB20 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 7.48Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

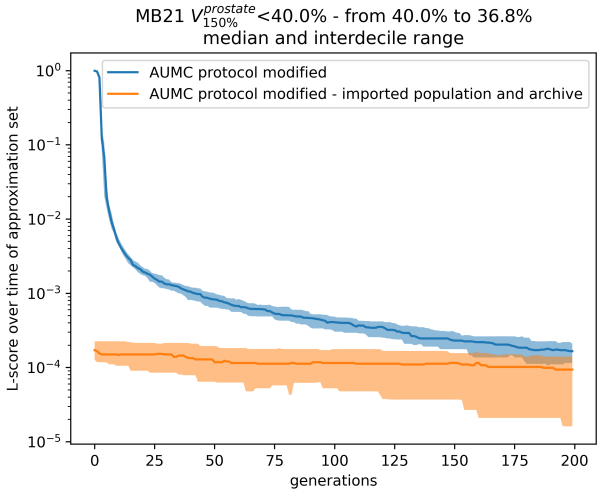

MB20 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 8.74Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

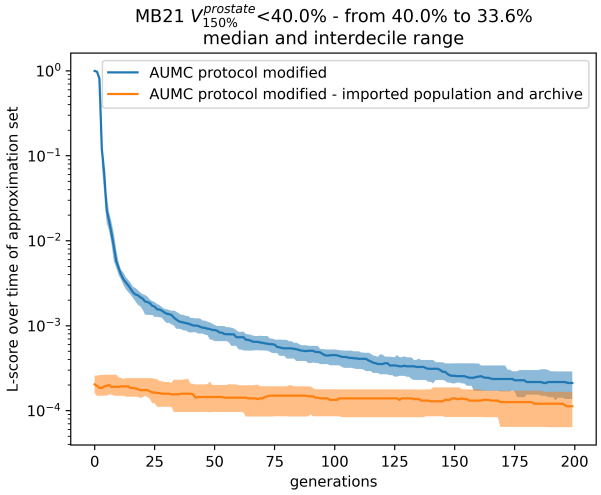

MB20 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 7.98Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

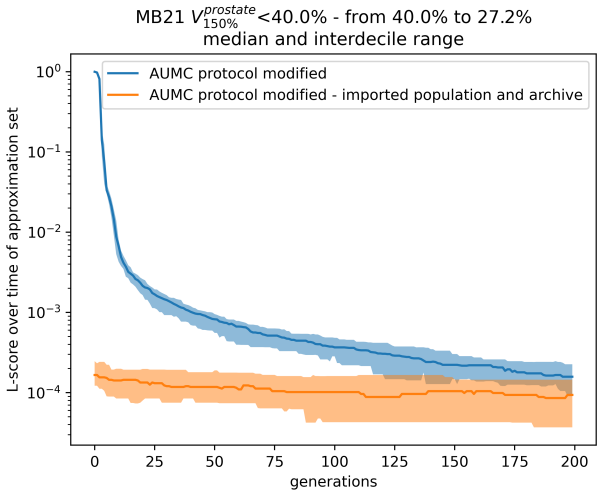

MB20 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 6.46Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

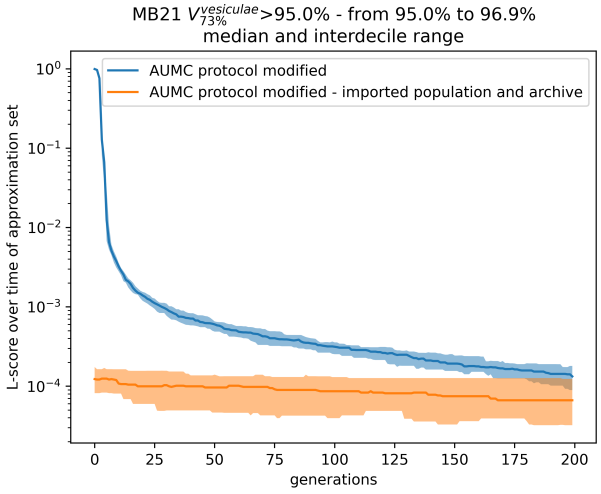

MB21 $V_{100\%}^{prostate}$>95.0% - from 95.0% to 96.9%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

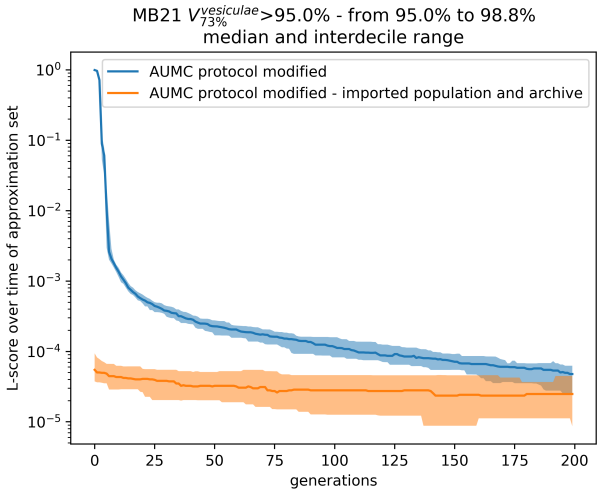

MB21 $V_{100\%}^{prostate}$>95.0% - from 95.0% to 98.8%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

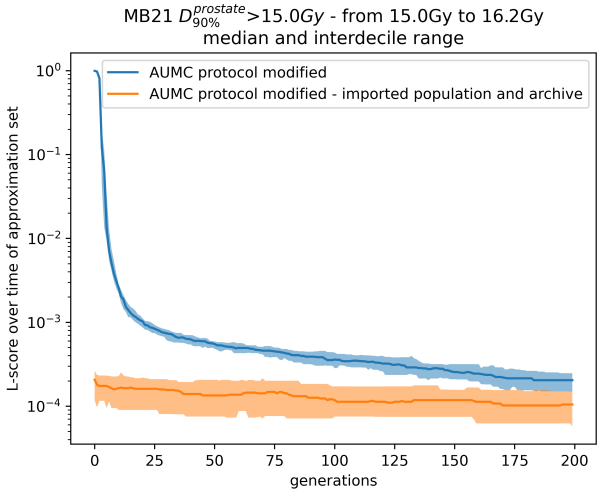

MB21 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 36.8%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

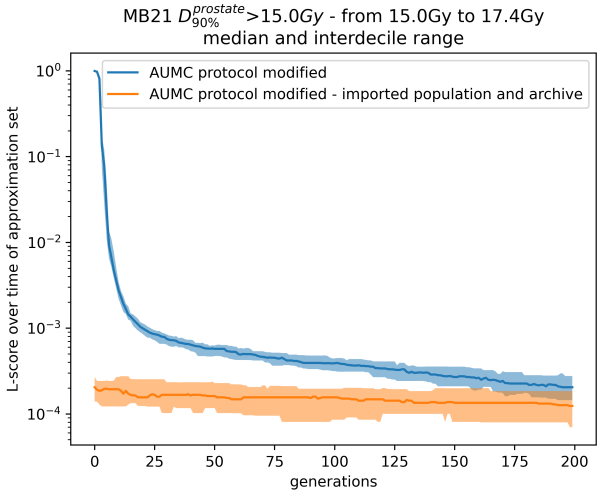

MB21 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 33.6%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

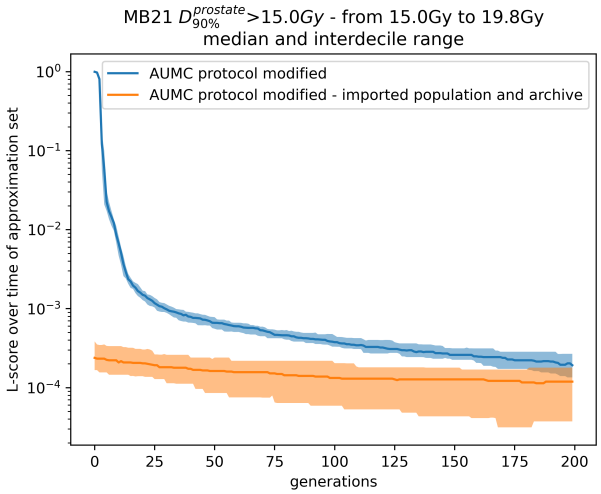

MB21 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 27.2%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

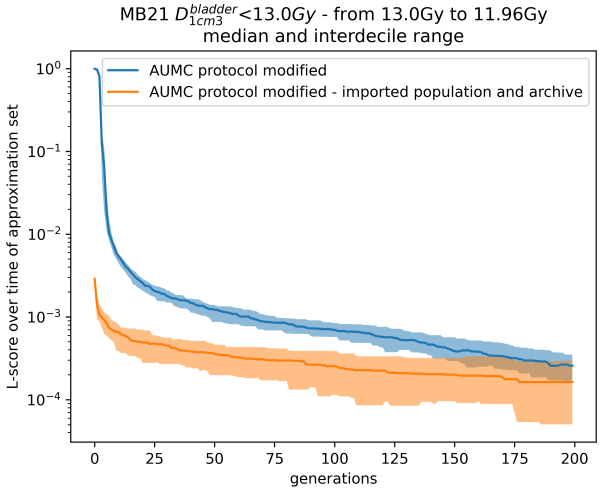

MB21 $V_{73\%}^{vesiculae}$>95.0% - from 95.0% to 96.9%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

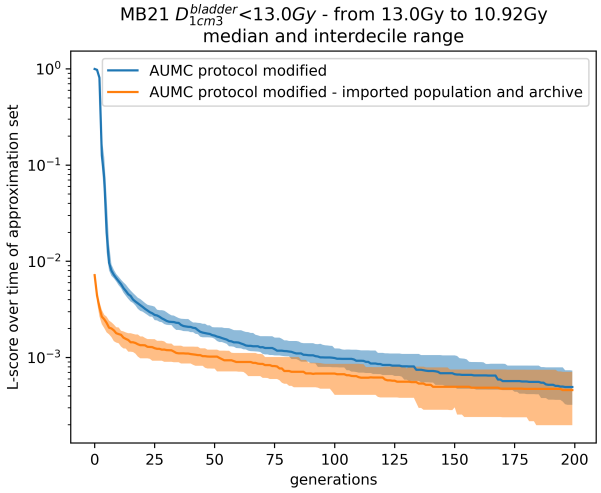

MB21 $V_{73\%}^{vesiculae}$>95.0% - from 95.0% to 98.8%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

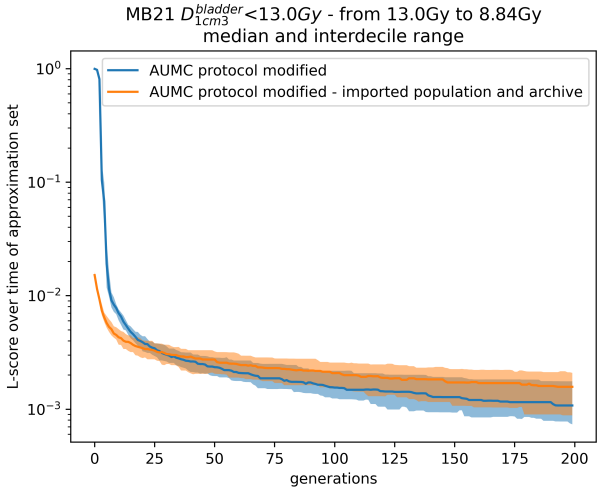

MB21 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 16.2Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

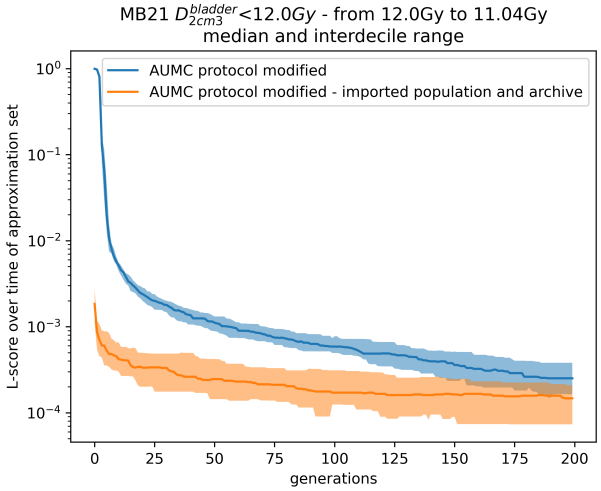

MB21 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 17.4Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

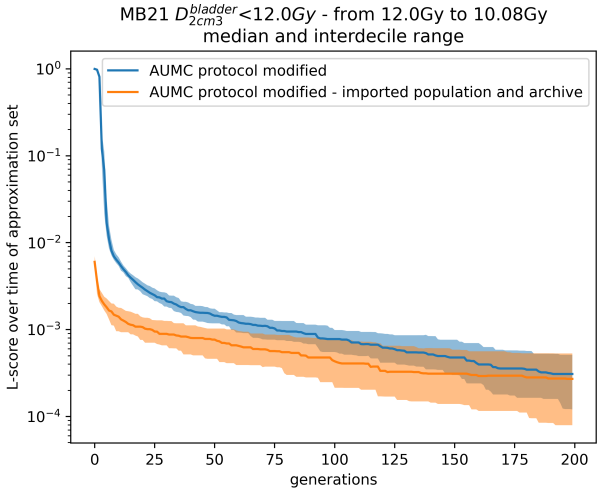

MB21 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 19.8Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

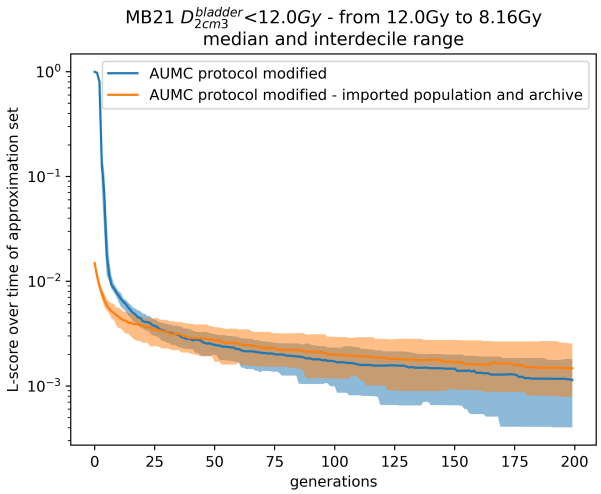

MB21 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 11.96Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

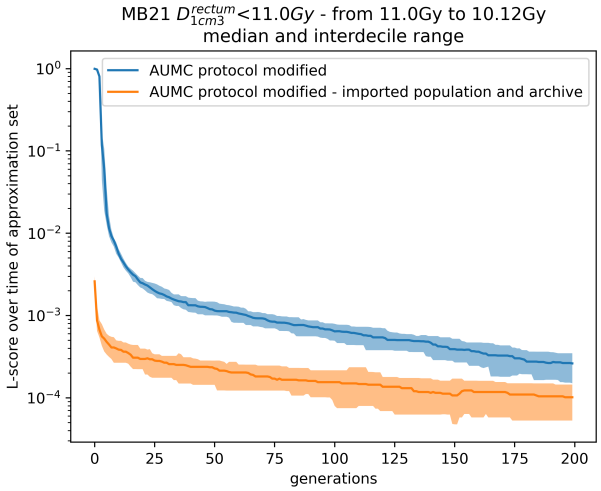

MB21 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 10.92Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

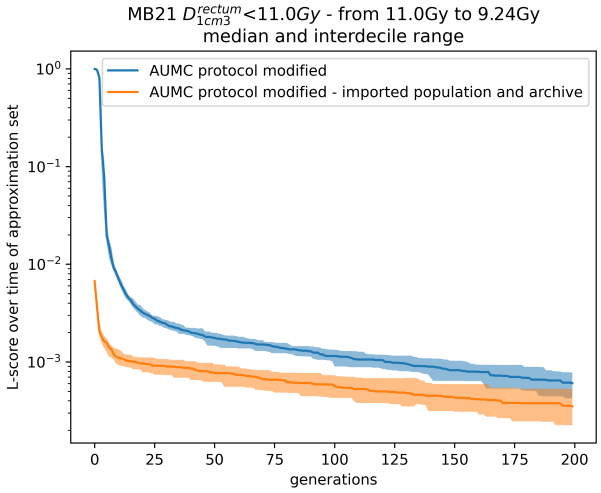

MB21 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 8.84Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

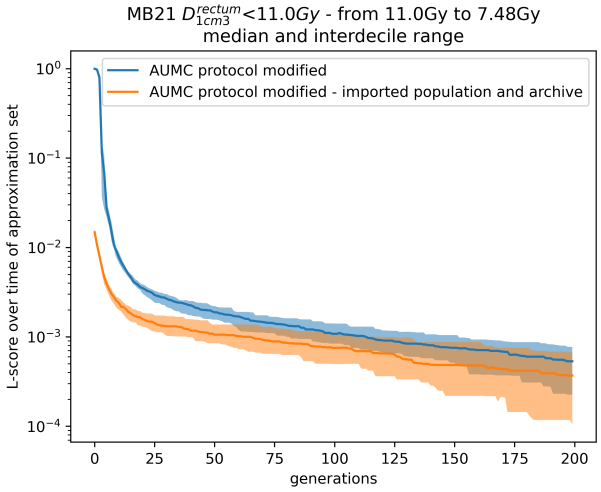

MB21 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 11.04Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

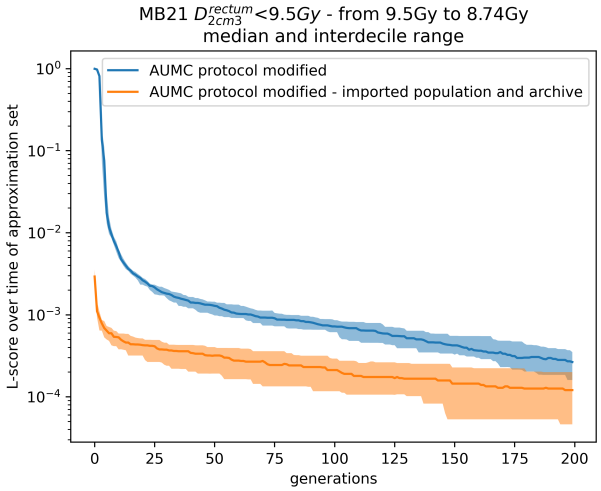

MB21 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 10.08Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

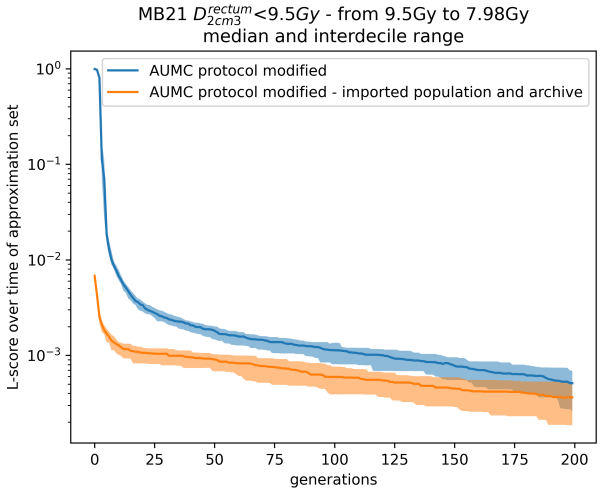

MB21 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 8.16Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

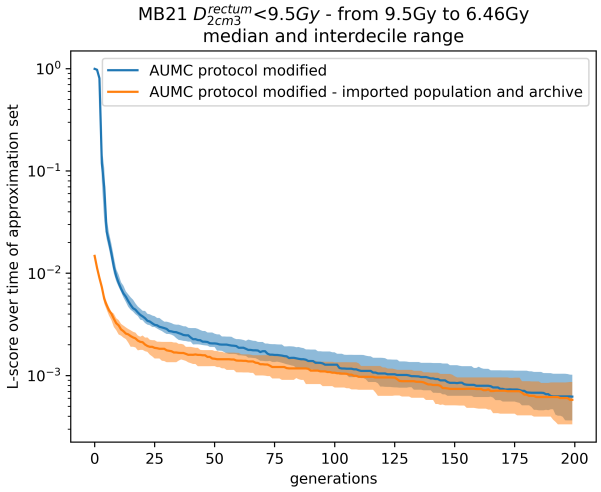

MB21 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 10.12Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

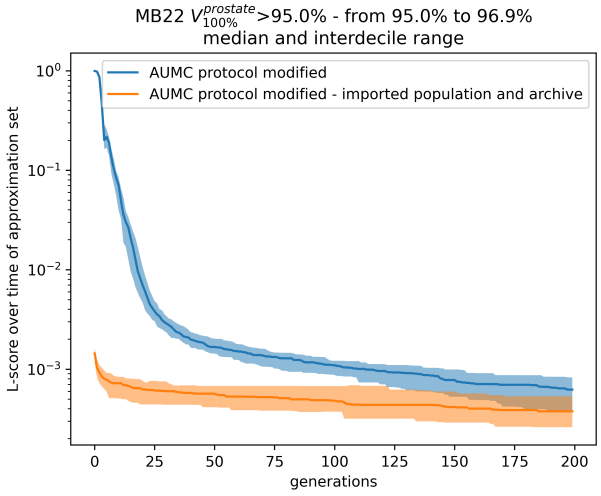

MB21 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 9.24Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

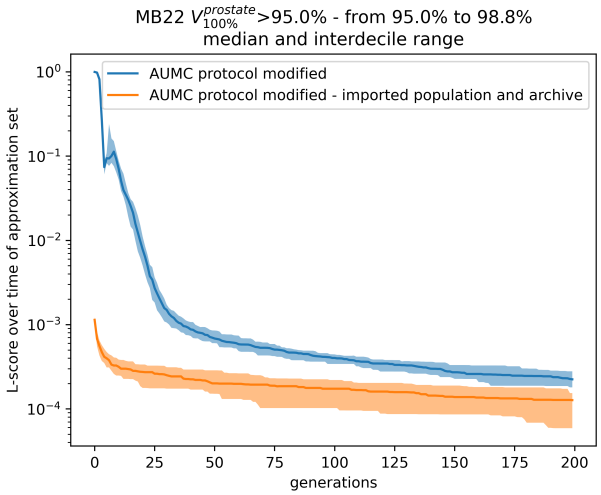

MB21 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 7.48Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

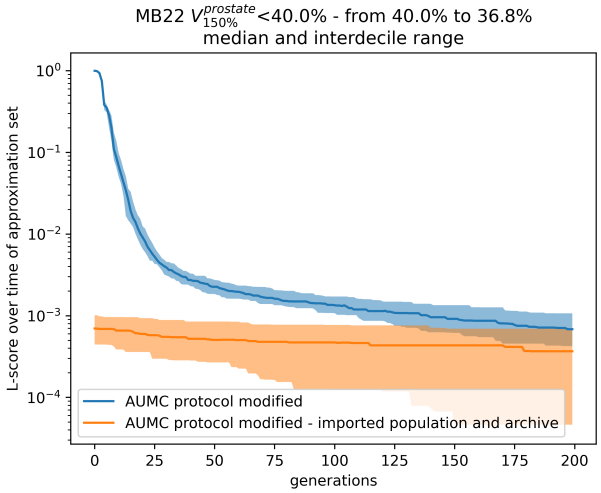

MB21 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 8.74Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

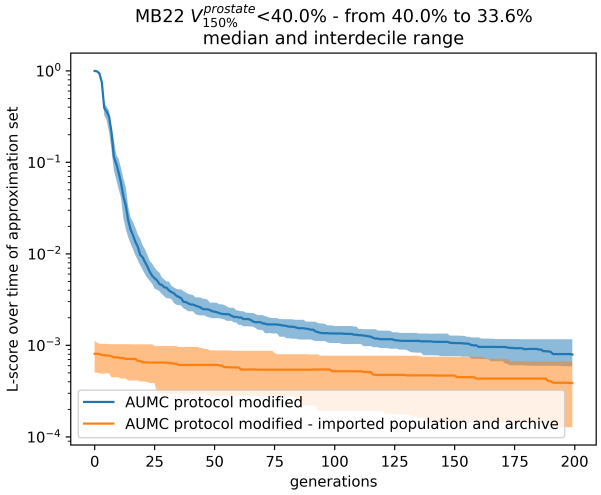

MB21 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 7.98Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

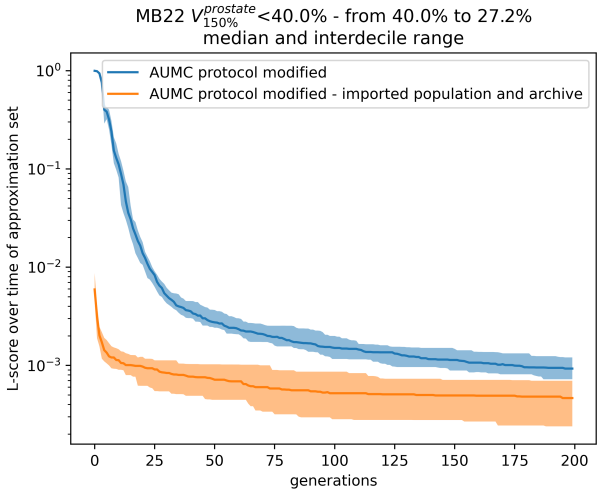

MB21 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 6.46Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

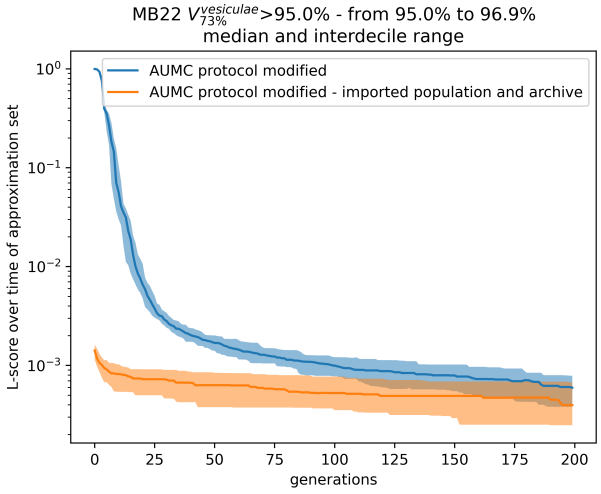

MB22 $V_{100\%}^{prostate}$>95.0% - from 95.0% to 96.9%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

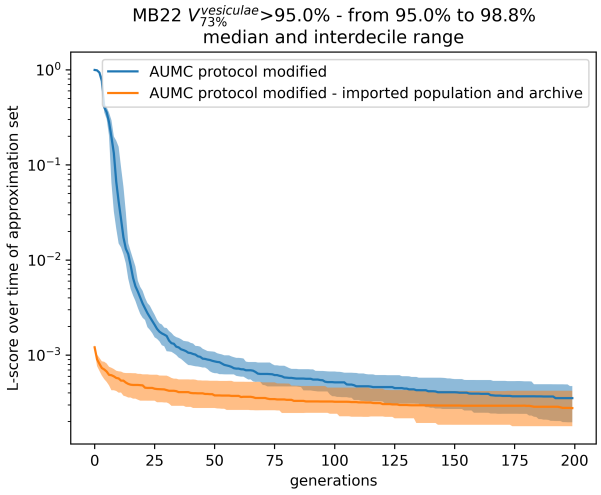

MB22 $V_{100\%}^{prostate}$>95.0% - from 95.0% to 98.8%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

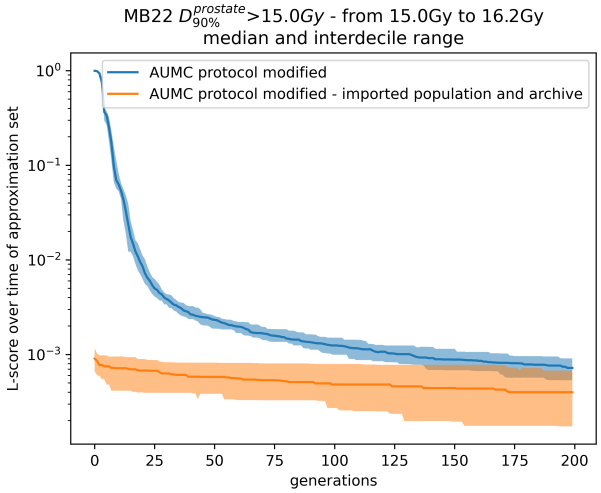

MB22 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 36.8%
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

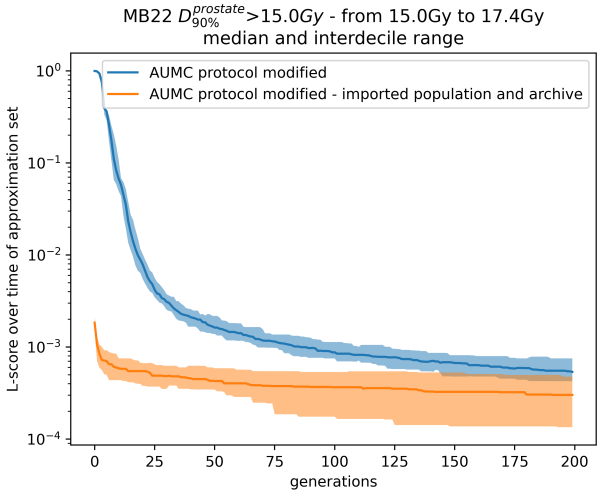

MB22 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 33.6%
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

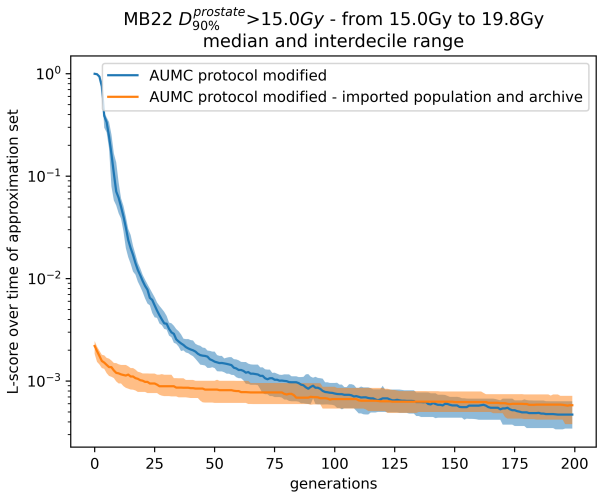

MB22 $V_{150\%}^{prostate}$<40.0% - from 40.0% to 27.2%
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

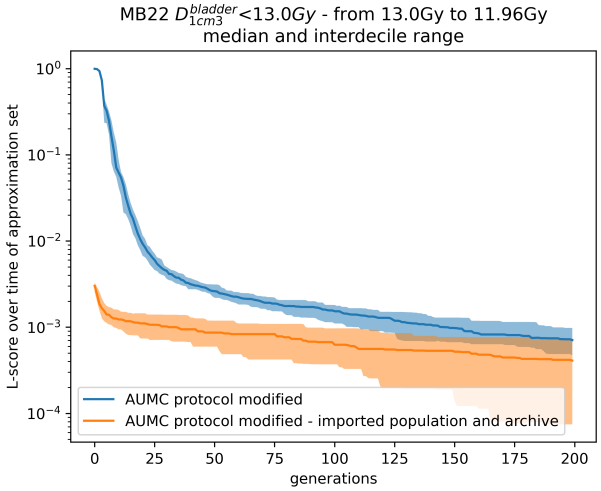

MB22 $V_{73\%}^{vesiculae}$>95.0% - from 95.0% to 96.9%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
10^{0}
10^{-1}
10^{-2}
10^{-3}
0
25
50
75
100
125
150
175
200
generations

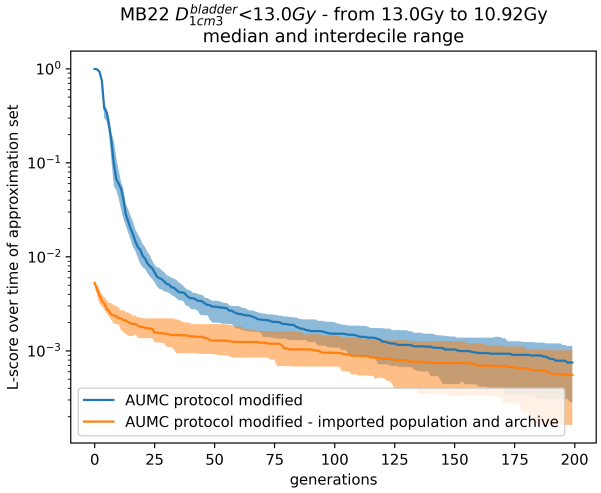

MB22 $V_{73\%}^{vesiculae}$>95.0% - from 95.0% to 98.8%
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
10^{0}
10^{-1}
10^{-2}
10^{-3}
0
25
50
75
100
125
150
175
200
generations

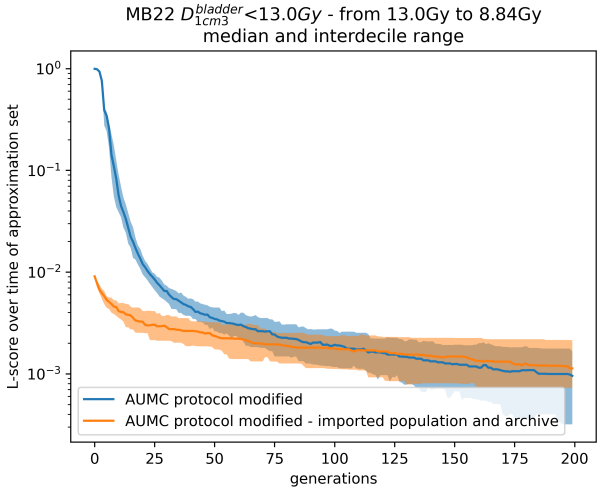

MB22 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 16.2Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

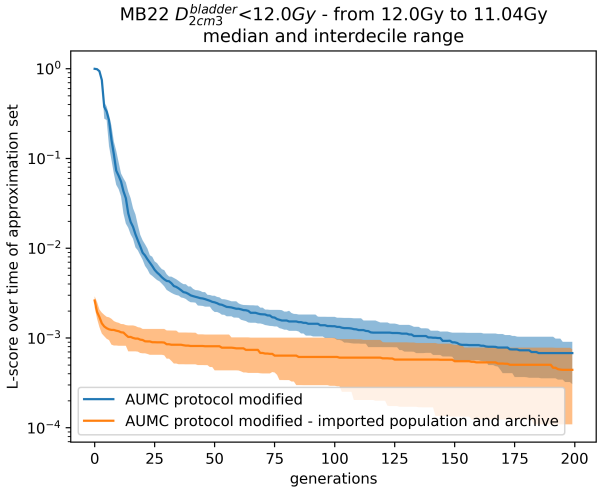

MB22 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 17.4Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

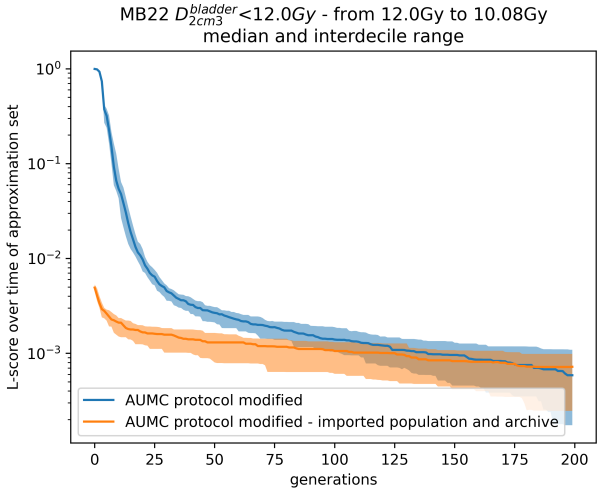

MB22 $D_{90\%}^{prostate}$>15.0Gy - from 15.0Gy to 19.8Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

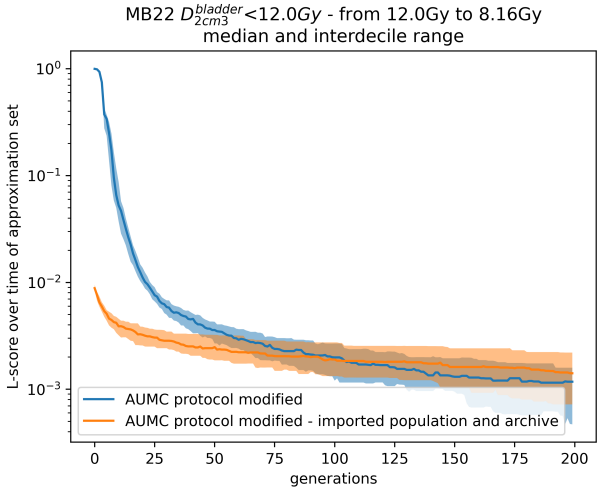

MB22 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 11.96Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

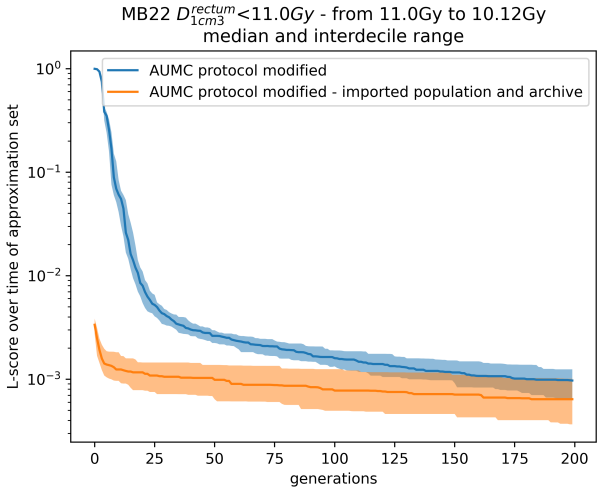

MB22 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 10.92Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

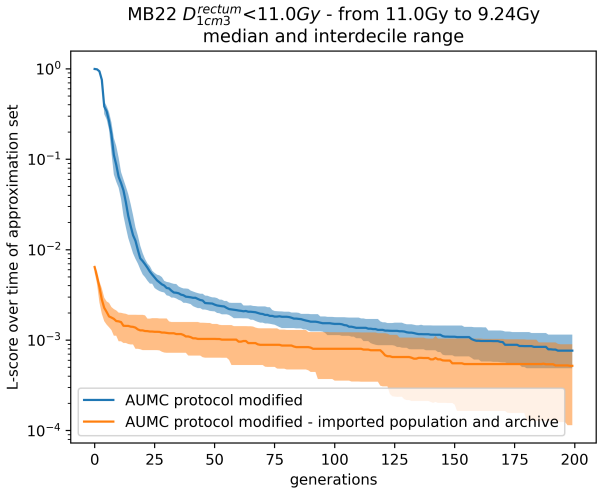

MB22 $D_{1cm3}^{bladder}$<13.0Gy - from 13.0Gy to 8.84Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

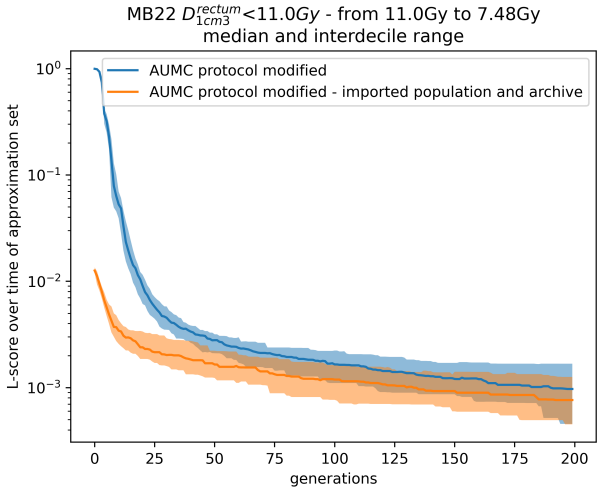

MB22 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 11.04Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

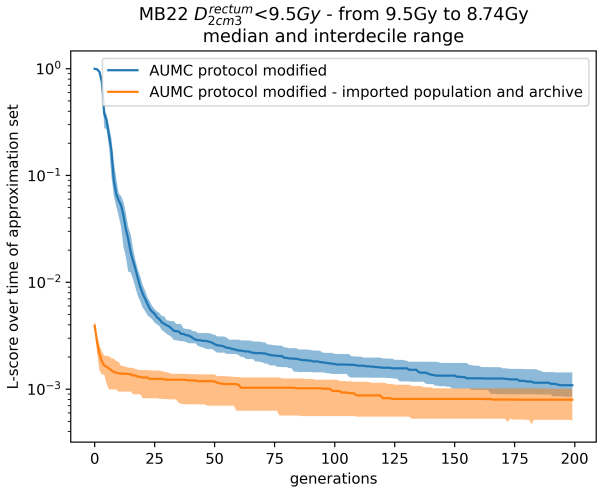

MB22 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 10.08Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

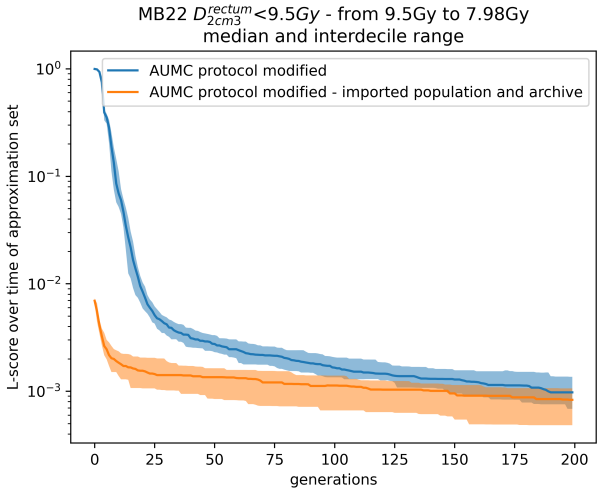

MB22 $D_{2cm3}^{bladder}$<12.0Gy - from 12.0Gy to 8.16Gy
median and interdecile range
L-score over time of approximation set
generations
AUMC protocol modified
AUMC protocol modified - imported population and archive

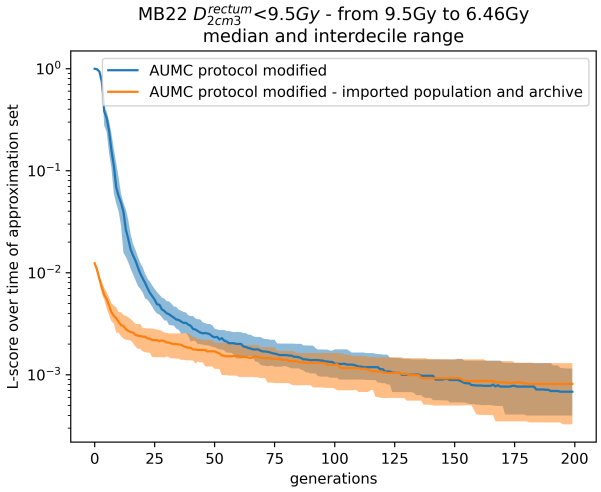

MB22 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 10.12Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations


MB22 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 9.24Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations


MB22 D_{1cm3}^{rectum}<11.0Gy - from 11.0Gy to 7.48Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations


MB22 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 8.74Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations


MB22 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 7.98Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations


MB22 D_{2cm3}^{rectum}<9.5Gy - from 9.5Gy to 6.46Gy
median and interdecile range
AUMC protocol modified
AUMC protocol modified - imported population and archive
L-score over time of approximation set
generations

Bibliography

Bouter, A., Alderliesten, T., Pieters, B. R., Bel, A., Niatsetski, Y., & Bosman, P. A. N. (2019). Gpu-accelerated bi-objective treatment planning for prostate high-dose-rate brachytherapy. Medical Physics, 46(9), 3776–3787. https://doi.org/https://doi.org/10.1002/mp.13681

Bouter, A., Alderliesten, T., Witteveen, C., & Bosman, P. A. N. (2017). Exploiting linkage information in real-valued optimization with the real-valued gene-pool optimal mixing evolutionary algorithm. Proceedings of the Genetic and Evolutionary Computation Conference, 705–712. https://doi.org/10.1145/3071178.3071272

Bouter, A., Luong, N., Witteveen, C., Alderliesten, T., & Bosman, P. (2017). The multi-objective real-valued gene-pool optimal mixing evolutionary algorithm [GECCO 2017 : Genetic and Evolutionary Computation Conference ; Conference date: 15-07-2017 Through 19-07-2017]. GECCO 2017 - Proceedings of the 2017 Genetic and Evolutionary Computation Conference, 537–544. https://doi.org/10.1145/3071178.3071274

Crook, J., Marbán, M., & Batchelar, D. (2020). Hdr prostate brachytherapy [Advances in Brachytherapy]. Seminars in Radiation Oncology, 30(1), 49–60. https://doi.org/https://doi.org/10.1016/j.semradonc.2019.08.003

Deb, K., & Deb, K. (2014). Multi-objective optimization. In E. K. Burke & G. Kendall (Eds.), Search methodologies: Introductory tutorials in optimization and decision support techniques (pp. 403–449). Springer US. https://doi.org/10.1007/978-1-4614-6940-7_15

Henry, A., Pieters, B. R., André Siebert, F., & Hoskin, P. (2022). Gec-estro acrop prostate brachytherapy guidelines. Radiotherapy and Oncology, 167, 244–251. https://doi.org/https://doi.org/10.1016/j.radonc.2021.12.047

Holland, J. H. (1992). Genetic algorithms. Scientific American, 267(1), 66–73. Retrieved June 13, 2022, from http://www.jstor.org/stable/24939139

Luong, H. N., & Bosman, P. A. N. (2012). Elitist archiving for multi-objective evolutionary algorithms: To adapt or not to adapt. In C. A. C. Coello, V. Cutello, K. Deb, S. Forrest, G. Nicosia, & M. Pavone (Eds.), Parallel problem solving from nature - ppsn xii (pp. 72–81). Springer Berlin Heidelberg.

Luong, N. H., Bouter, A., van der Meer, M. C., Niatsetski, Y., Witteveen, C., Bel, A., Alderliesten, T., & Bosman, P. A. N. (2017). Efficient, effective, and insightful tackling of the high-dose-rate brachytherapy treatment planning problem for prostate cancer using evolutionary multi-objective optimization algorithms. Proceedings of the Genetic and Evolutionary Computation Conference Companion, 1372–1379. https://doi.org/10.1145/3067695.3082491

Luong, N. H., La Poutré, H., & Bosman, P. A. (2014). Multi-objective gene-pool optimal mixing evolutionary algorithms. Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation, 357–364. https://doi.org/10.1145/2576768.2598261

Maree, S. C., Luong, N. H., Kooreman, E. S., van Wieringen, N., Bel, A., Hinnen, K. A., Westerveld, H., Pieters, B. R., Bosman, P. A., & Alderliesten, T. (2019). Evaluation of bi-objective treatment planning for high-dose-rate prostate brachytherapy—a retrospective observer study. Brachytherapy, 18(3), 396–403. https://doi.org/https://doi.org/10.1016/j.brachy.2018.12.010

Nazar, M., Alam, M. M., Yafi, E., & Su'ud, M. M. (2021). A systematic review of human–computer interaction and explainable artificial intelligence in healthcare with artificial intelligence techniques. IEEE Access, 9, 153316–153348. https://doi.org/10.1109/ACCESS.2021.3127881

Porter, A., Blasko, J., Grimm, P., Reddy, S., & Ragde, H. (1995). Brachytherapy for prostate cancer. CA: a cancer journal for clinicians, 45(3), 165–178. https://doi.org/10.3322/canjclin.45.3.165

Thierens, D., & Bosman, P. (2011). Optimal mixing evolutionary algorithms. IEEE/ACM Transactions on Computational Biology and Bioinformatics - TCBB, 617–624. https://doi.org/10.1145/2001576.2001661

Thierens, D., & Bosman, P. A. (2011). Optimal mixing evolutionary algorithms. Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation, 617–624. https://doi.org/10.1145/2001576.2001661

van der Meer, M. C., Bosman, P. A. N., Pieters, B. R., Niatsetski, Y., van Wieringen, N., Alderliesten, T., & Bel, A. (2019). Sensitivity of dose-volume indices to computation settings in high-dose-rate prostate brachytherapy treatment plan evaluation. Journal of Applied Clinical Medical Physics, 20(4), 66–74. https://doi.org/https://doi.org/10.1002/acm2.12563