# Classical Dosimetry Automation

## Contents




**Abstract**

In radiation therapy, treatment planning involves balancing competing objectives. The contradictory goals often lack universal prioritization. Expert bias introduces variability in clinical practice, as the preferences of radiation oncologists and medical physicists shape treatment planning. Traditionally, this balance is achieved through manual or semi-manual processes guided by the expertise of clinicians and planners. This chapter explores approaches for fully automating the treatment planning process, focusing on classical optimization techniques constrained by dosimetric objectives.

We will review established methods and propose new agents capable of optimizing dose without human interaction. This innovative approach leverages previously defined dose distance metrics. We aim to streamline and standardize the treatment planning workflow by fully automating the optimization process.


## 1 Meta Optimization Approach

After the FMO step is developed and encapsulated, one may modify the weights used for the FMO iteratively until a condition is met or for a given number of steps. We will, therefore, have one inner optimization (the FMO) and one outer optimization.

**Algorithm 1** Meta Optimization Algorithm Outline

```
initialize w
repeat
    initialize b                              ▷ FMO starts
    repeat
        d = Lb                                ▷ differentiable
        c = C(w, d)                           ▷ differentiable
        back-propagate c
        update b
    until FMO stop condition                  ▷ FMO ends
    update w
until Meta-optimization stop condition
```

The outer optimization step is not differentiable (or at least not in a reasonable computation time). Hence, we will be looking at gradient-free optimization methods.

## 1.1 Expert Weight Adjustment

Expert systems are computer systems emulating the decision-making of a human expert.

**Simple Weight Increase** One approach involves increasing the weight of all unsatisfied constraints after each FMO optimization step. This method is advantageous due to its simplicity in terms of implementation and understanding. However, a significant limitation arises when none of the constraints are met, causing the outer optimization loop to stagnate. In such cases, the optimization process remains stationary, usually when too many constraints are enforced. This stationary state arises particularly in complex scenarios with multiple competing constraints and can result in a situation where progress is hindered, preventing the solution from improving over iterations.

**Inverse Proportional Weight Increase** Another approach involves increasing the weight of each constraint inversely proportional to how close it is to being met, thereby quantifying the degree of constraint satisfaction. For instance, the degree of satisfaction can be quantified by calculating the area between the dose-volume histogram (DVH) constraint and the actual DVH curve; when this area is zero, the constraint is considered fully satisfied. While this method remains relatively straightforward to implement and provides a more refined adjustment of weights based on how close each constraint is to being met, it can lead to oscillation issues. Constraints may fluctuate between being satisfied and violated across iterations, hindering stable convergence. Although adding momentum to the optimization process could mitigate these oscillations, expert systems of this nature typically require continuous tuning and refinement. As a result, this approach may not be viable for reliable clinical applications where consistent performance is essential.

## 1.2 Metric-Based

Here, we suppose that one can construct a measure of the quality of a plan.

**Hill Climbing** Hill climbing [Ski08] is a simple optimization technique in which the solution iteratively moves toward an improved solution based on a defined metric. In the context of radiotherapy treatment planning, several metrics have been proposed to quantify the quality of a plan, including Normal Tissue Complication Probabilities (NTCP), target coverage, conformity index, and heterogeneity index, among others [Lym92, LGW+22] . This approach offers a systematic way to improve treatment plans by optimizing the chosen metric.

However, defining the correct metric of interest is challenging, as no single metric, nor a combination of metrics, has consistently proven to satisfy radio-oncologists' requirements. In practice, the most reliable method for assessing the quality of a treatment plan remains the manual evaluation of dose-volume histograms (DVHs), which provide a detailed representation of the dose distribution across both the target and the surrounding organs at risk.

**Pareto Exploration** Researchers have developed algorithms to explore the Pareto surface of dose distributions, yet no consensus has been reached on selecting an optimal dose from this surface. Consequently, Pareto surface exploration is unsuitable due to the absence of an objective quantitative measure for evaluating the quality of a specific plan [HYP+21]. This limitation similarly constrains other meta-optimization techniques, as they also rely on the availability of a clear, impartial criterion for plan evaluation [WZ01, XLD+99].

**Contextual Knowledge** Another challenge is the varying difficulty across patients due to their different organ geometry. In "easy" cases, clinicians may require a highly optimized dose distribution regarding the previously mentioned metrics. On the other hand, for "harder" cases, they can afford to be more lenient regarding constraint satisfaction.

This context-aware acceptability criterion adds complexity to the optimization process. It becomes challenging to define general rules not only for ranking treatment plans but also for deciding a plan's acceptability.

# 2 Radiotherapy Dose Optimization via Clinical Knowledge Based Reinforcement Learning (AIME 2024)

## 2.1 Introduction

Reinforcement learning (RL) is a machine learning paradigm that trains agents to make sequential decisions in dynamic environments [Bro21]. Agents learn to optimize their actions to achieve long-term objectives through trial and error guided by rewards or penalties. The decisions taken by dosimetrists when optimizing treatment can be formalized as an RL problem. Moreover, dosimetrists can guide the TPS towards an acceptable plan but usually struggle to explain their decision while interacting with the TPS. The difficulty in explaining why certain decisions are taken suggests using deep RL over expert-based methods. This setup is similar to image recognition, where one can say a picture represents a car or a boat but struggles to explain why.

The study's primary hypothesis is that all the information needed to decide what weights should be changed in the objective function used by the optimizer relies on the Dose Volume Histograms (DVHs). Our hypothesis is supported by the fact that dosimetrists almost solely use DVHs plots. In order to learn the actions of dosimetrists who use a TPS to optimize doses, we leverage deep learning. This is done by training an agent that takes the DVHs as the input of the current optimized dose, and predicts the evaluation of possible weights changes.

Typically, access to the exact actions taken by human dosimetrists on the TPS is unavailable (as clinics do not usually store this data; only the final plan is held). Therefore, we only use the dose distributions of previously treated patients to train our model. This partial availability of data suggests the use of RL.

### 2.2 Materials and Methods

#### 2.2.1 Reinforcement Learning Reward

#### 2.2.2 Architecture

#### 2.2.3 Avoiding Off-Distribution

#### 2.2.4 Quantitative Results

#### 2.2.5 Qualitative Results

### 2.3 Discussion

### 2.4 Conclusion

### Appendix

#### Synthetic phantom patients

#### Clinical dose

#### Optimization

## 3 Clinically Dependent Fully Automatic Treatment Planning System (ASTRO 2024)

### 3.1 Purpose / Objective

### 3.2 Materials/Methods

### 3.3 Results

### 3.4 Conclusion

## References


[Bro21] Ruth Brooks. What is reinforcement learning?, December 2021.

[HYP+21] Charles Huang, Yong Yang, Neil Panjwani, Stephen Boyd, and Lei Xing. Pareto Optimal Projection Search (POPS): Automated Radiation Therapy Treatment Planning by Direct Search of the Pareto Surface. *IEEE Transactions on Biomedical Engineering*, 68(10):2907–2917, October 2021.

[LGW+22] Xinyi Li, Yaorong Ge, Qiuwen Wu, Chunhao Wang, Yang Sheng, Wentao Wang, Hunter Stephens, Fang-Fang Yin, and Q. Jackie Wu. Input feature design and its impact on the performance of deep learning models for predicting fluence maps in intensity-modulated radiation therapy. *Physics in Medicine & Biology*, 67(21):215009, November 2022.

[Lym92] John T. Lyman. Normal tissue complication probabilities: Variable dose per fraction. *International Journal of Radiation Oncology*Biology*Physics*, 22(2):247–250, January 1992.

[Ski08] Steven S. Skiena. *The Algorithm Design Manual*. Springer, London, 2008.

[WZ01] Xingen Wu and Yunping Zhu. An optimization method for importance factors and beam weights based on genetic algorithms for radiotherapy treatment planning. *Physics in Medicine and Biology*, 46(4):1085–1099, April 2001.

[XLD+99] L Xing, J G Li, S Donaldson, Q T Le, and A L Boyer. Optimization of importance factors in inverse planning. *Physics in Medicine and Biology*, 44(10):2525–2536, October 1999.