Technical Innovations & Patient Support in Radiation Oncology 37 (2026) 100383

Contents lists available at ScienceDirect

Technical Innovations & Patient Support in Radiation Oncology

journal homepage: www.sciencedirect.com/journal/technical-innovations-and-patient-support-in-radiation-oncology

Simulation and optimization of treatment schedule for multi-gantry heavy ion therapy

Ethan Stolen[a,1], Sridhar Yaddanapudi[b,1,*], Helen Qian[a], Sang Kyun Yoo[a,c,d], Bo Lu[b], Justin C. Park[b], Jun Tan[b], Xiaoying Liang[b], Keith M. Furutani[b], Chris J. Beltran[b], James J. Sohn[a,*]

[a] *Department of Radiation and Cellular Oncology, University of Chicago, Chicago, IL 60611, USA*
[b] *Department of Radiation Oncology, Mayo Clinic Florida, Jacksonville, FL 32224, USA*
[c] *Yonsei Institute for Digital Health, Yonsei University, Seoul 03722, South Korea*
[d] *Department of Radiation Oncology, Yonsei Cancer Center, Heavy Ion Therapy Research Institute, Yonsei University College of Medicine, Seoul 03722, South Korea*

ARTICLE INFO

Keywords:
Carbon ion radiotherapy
Heavy ion therapy
Schedule optimization
Bayesian optimization

ABSTRACT

Background: Carbon ion radiotherapy (CIRT) facilities face scheduling challenges due to multiple treatment rooms sharing a single synchrotron, leading to extended patient wait times and reduced efficiency. This study developed and evaluated an optimization tool for patient scheduling in multi-gantry CIRT facilities to minimize in-room patient wait time and maximize patient throughput.
Methods: A simulation–optimization model was created to emulate patient flow in a facility with one fixed-beam and three gantry rooms. We developed a genetic algorithm (GA) to balance competing objectives like wait times, room utilization, and treatment priority. Its performance was evaluated against baseline scheduling and a Bayesian optimization (BO) approach, using metrics of average in-room wait time and algorithm execution speed. Robustness was tested across various scenarios, including equipment downtime.
Results: The GA reduced average patient wait times by 17% (from 0.47 to 0.39 min) compared to baseline. BO performed significantly better, achieving a 92% reduction (to 0.04 min). While GA converged faster (3 min), BO required more iterations to converge, thereby taking longer (7 min), but achieving much greater reduction in average patient wait times. In 100 independent test scenarios, the BO approach significantly reduced wait times in 100% of the simulations, compared to 19% for the GA. Both algorithms maintained robust performance across diverse conditions, demonstrating their reliability. Analysis of outliers ($N = 100$) showed that BO significantly reduced the maximum patient wait time from 3.52 to 1.33 min ($p < 0.001$), minimizing clinically relevant extremes.
Conclusion: This application of scheduling optimization to the unique constraints of CIRT shows potential for improving efficiency at CIRT facilities. While requiring minutes rather than seconds to converge, the BO algorithm effectively eliminates immobilized wait times, making it a powerful tool for daily schedule optimization. Future research will focus on validation with empirical data from operational CIRT centers.

Introduction

The therapeutic use of high-energy protons was first proposed in 1946 by Robert Wilson [1]. The first patient treatment using protons occurred at Lawrence Berkeley Laboratory in 1954 [2]. Since then, the use of heavy ions, such as carbon, for cancer treatment has increased significantly. According to the Particle Therapy Co-Operative Group (PTCOG) [3], more than 17 facilities across Asia and Europe offer carbon-ion radiation therapy (CIRT), and efforts are underway to establish the first CIRT center in the United States [4,5]. Particle therapy, using protons and carbon ions, is known for its precise energy deposition and allows for highly localized dose delivery, minimizing lateral scattering in deeply seated tissues. As a result, particle therapy offers significant advantages over traditional X-ray therapy. Clinical

* Corresponding authors.
E-mail addresses: yaddanapudi.sridhar@mayo.edu (S. Yaddanapudi), jooyoung.sohn@bsd.uchicago.edu (J.J. Sohn).
[1] These authors contributed equally.

https://doi.org/10.1016/j.tipsro.2026.100383
Received 8 December 2025; Received in revised form 4 February 2026; Accepted 6 February 2026
Available online 7 February 2026

studies have shown that CIRT improves treatment efficacy, leading to better local tumor control and enhanced overall survival rates in various types of cancer [6].

Heavy ion delivery systems use a synchrotron to accelerate ions to high energies. Magnets in the beam transport system then guide and focus these ions toward different treatment rooms. Typically, a center may have several treatment rooms, each capable of delivering treatment using a rotating gantry or a fixed-beam line. The gantry rooms offer multi-angle beam delivery and greater flexibility, while the fixed-beam line delivers the beam at a set angle, requiring patients to be repositioned for multi-angle treatments [7–9].

One of the main challenges that heavy ion treatment facilities face is efficiently managing patient flow. Unlike conventional radiotherapy, where each linear accelerator operates independently, particle therapy relies on a single synchrotron for beam generation. This shared resource means that only one patient can receive treatment at a time, leading to longer wait times for patients. These delays can lead to discomfort and lower patient satisfaction. In this study, wait time is defined as the duration from the end of patient setup to the start of beam delivery. From a patient perspective, this period is particularly challenging as individuals remain fully immobilized in masks, thermoplastic shells, or other devices designed to ensure sub-millimeter positioning accuracy. Even brief delays can cause significant physical discomfort, anxiety, and claustrophobia [10–12]. Additionally, prolonged wait times result in suboptimal use of the beam, decreasing patient throughput in the facility. Therefore, reducing patient wait times and optimizing scheduling is beneficial for enhancing the treatment experience and improving efficiency.

Heavy ion treatment facilities employ several scheduling and optimization strategies to enhance efficiency and minimize patient delays. Optimizing for patient wait time poses a few challenges in a heavy ion treatment facility. The first is the existence of multiple treatment rooms competing for the beam. There is also significant variability in the time required for patient setup, beam-on time, and the number of beams a particular patient requires. This is further complicated by the fixed-beam line room, which requires a different and longer setup procedure compared to the gantry room. Recent studies have proposed several optimization algorithms to address the issue of patient wait time. The gatekeeper approach [13] identifies patients with long wait times and adjusts the callback for a particular patient accordingly. The simulation optimization gantry call-back control (SOGCC) [14] and rollout-based gantry call-back control (RGCC) [15] methods both allow for real-time decision making based on simulated patient scheduling. However, none of these methods can optimize for extraneous factors not defined by the simulation. For example, patient arrival rate, mean arrival time, and the probability that a patient may need the fixed-beam line room all affect patient wait time but are not adequately addressed by these algorithms. Additionally, none of these algorithms has been implemented in heavy ion facilities, which pose challenges to patient scheduling. In this proof of concept study, we apply genetic algorithm (GA) and Bayesian optimization (BO) approaches to optimize a detailed simulation model of a CIRT facility. The gatekeeper approach is a heuristic strategy that prioritizes patients based on predefined rules, including identifying those with long accumulated wait times and adjusting their call-back priority to prevent excessive delays. Unlike this reactive, rule-based method, the GA and BO approaches proposed here are predictive and stochastic, optimizing the entire day's schedule structure rather than managing individual patient queues dynamically. Our goal is to demonstrate a practical method for reducing this immobilized wait time, thereby improving patient comfort and increasing overall patient throughput by optimizing resource utilization. This work establishes the foundation for subsequent validation studies using empirical data from operational CIRT centers.

Methods

Model overview

This paper integrates simulation and optimization techniques in order to address the issue of patient wait time within the treatment room in a simulation–optimization model, developed using the open-source SimPy package [16] in Python 3.12. The simulation portion of our model consisted of all events in the workflow for treating patients, emphasizing patient waiting times. In the simulation, each patient progressed through room entry time, setup time, beam waiting time, beam delivery time, mid-treatment setup event (i.e., the time required to reposition the patient, only in the fixed-beam room), exit time, and turnaround time. The simulated facility included three gantry rooms along with a single fixed-beam room. Patients were randomly generated with assignments to either a gantry or fixed-beam room, along with the required number of treatment fields and predefined setup and treatment times. These parameters were designed to reflect real-world caseload data [13,17]. The simulated facility models a mixed carbon ion and proton beam facility, which includes a patient population requiring carbon-ion-specific treatments alongside others who could be treated with either modality.

In our simulation, each patient was randomly assigned between 1–4 treatment beams (fields) based on realistic treatment planning requirements. This variability in beam count significantly impacts scheduling complexity, as patients requiring multiple fields need longer overall treatment times and place additional demands on the shared beam resource. Importantly, our model incorporates a critical operational constraint: patients assigned to the gantry rooms must be treated in that room type due to their specific treatment requirements, including multiple beam angles, while patients assigned to the fixed-beam room maintain flexibility and can be reassigned between different gantry treatment rooms and the fixed-beam room during the pre-arrival scheduling phase. This room assignment flexibility for fixed-beam room patients represents an important degree of freedom that the optimization algorithms can leverage to improve overall scheduling efficiency.

Patient call-back occurs at the beginning of event e, where entry time is determined through a Monte Carlo approach between values generally accepted by current literature, as shown in Table 1. This same approach is applied to S_f, S_g, T_f, T_g, M_f, and n. Patients are pre-assigned during the treatment planning process to either the fixed-beam room or one of the

Table 1
Description of parameters used in optimization study.

Parameter	Description	Typical Value	Reference
P_g	Set of patients to be scheduled on the gantry	80–100 Patients	[13]
P_f	Set of patients to be scheduled on fixed-beam	5–15 Patients	
n	Number of fields required for a particular patient	1–4 fields	[17]
w	Wait time, end of setup to start of treatment	5–15 min.	
T_g	Treatment time for one gantry beam	1–2 min.	[17]
T_f	Treatment time for one fixed-beam	1–2 min.	[17]
B	Total beam on time	1–25 min.	[17]
e	Entry time	2–5 min.	
S_g	Setup time for gantry	5–20 min.	[13,17]
S_f	Setup time for fixed-beam	15–30 min.	
M_f	Mid-treatment setup time in fixed-beam room	10–15 min.	
E	Exit time	1–2 min.	[17]
t	Room turnaround time	5–10 min.	

gantry rooms based on their specific tumor type and treatment requirements. When a patient checks in for treatments, they will be directed to their pre-assigned treatment room. Should the designated room not be available at the scheduled time, the patient will remain in the queue in the waiting room until their assigned room becomes available. In either of the gantry rooms, a patient is set up and positioned in the gantry during S_g. This same logic applies to the fixed-beam room, with setup in the fixed-beam room designated as S_f. We make this distinction to accommodate the longer setup time. In an ideal scenario, treatment time, otherwise known as beam-on time, would follow immediately after the end of setup, denoted as T_g and T_f. A patient may have any number of treatment fields as designated by the patient generator function. If a patient requires multiple fields in the fixed-beam room, setup time between fields is designated as M_f, or mid-treatment setup time. The patient flow model concludes with patient exit time, denoted by E. This is followed by turnaround time, t, a set amount of time required to clean the room in preparation for the next patient. The total simulation time models a full 8-hour workday (480 min). Fig. 1 displays a representative 120-minute window of this schedule to ensure readability of the scheduling dynamics.

Between the end of setup and the beginning of beam-on time, or actual treatment time, is the wait time within the treatment room, designated as w. When the beam is in use in a room other than the room that requested it, the latter enters a queue for the beam. The queue operates on a first-in, first-out principle based on availability. In such a multiroom treatment facility, this conflict is inevitable but unfortunately results in considerable patient discomfort. This wait time, w, is the primary focus as it represents time the patient is immobilized and uncomfortable. These parameters (setup times, beam-on times, number of fields) were sampled from distributions defined by published operational data from active particle therapy centers [13,17] [Table 1], ensuring the simulation statistically reflects realistic clinical loads (Fig. 2).

Optimization algorithm

The primary objective of our optimization algorithm is to minimize the average immobilized patient wait time. The total wait time is determined by setup time, mid-treatment setup time, and treatment time. To optimize the average wait time, it is calculated by dividing the

Fig. 1. Randomly generated patient schedule, within parameter constraints, shown before and after the schedule is optimized using the genetic algorithm (GA). The average individual wait time before the optimization was 1.09 min, and the optimized average individual wait time was 0.25 min for the BO, and 0.76 min for the GA optimization. The example chosen represented a schedule with a very long initial wait time, which may increase the performance of the optimization algorithms compared to the average. Additionally, a high-load stress-test scenario with an initial average wait time of 1.09 min, significantly higher than the global average of 0.45 min reported in Table 2, was selected to visually illustrate the algorithms' capacity to resolve complex resource conflicts.

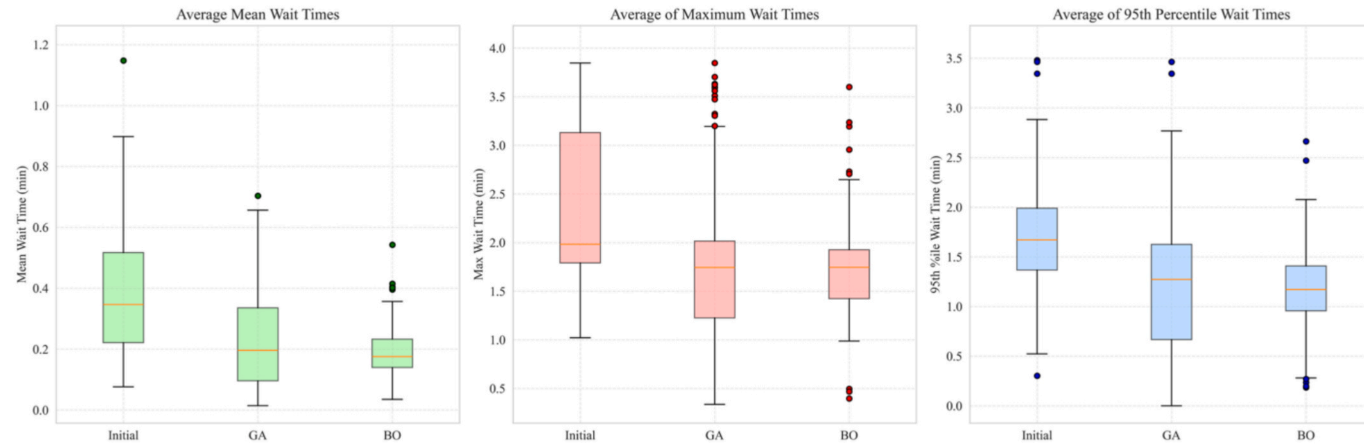

Fig. 2. Comparison of patient wait time metrics across 1,000 independent simulations. The box plots illustrate the distribution of Mean Wait Time, 95th Percentile Wait Time, and Maximum Wait Time for the Initial (Baseline) schedule, Genetic Algorithm (GA), and Bayesian Optimization (BO). The central line represents the median, the box limits indicate the interquartile range (IQR, 25th to 75th percentiles), and the whiskers extend to 1.5 × IQR. Points beyond the whiskers represent individual outliers. While all metrics show improvement, the reduction in Maximum Wait Time demonstrates the algorithms' ability to eliminate extreme outliers.

total wait time by the number of patients using both the gantry and fixed-beam treatment rooms, as represented by Eq. (1).

$$w_{avg} = \frac{\sum w}{(P_g + P_f)} \quad (1)$$

Our fitness function is designed to minimize the average wait time by taking the negative value of this equation. Two optimization techniques, genetic algorithm (GA) [18] and Bayesian optimization (BO) [19], were chosen for their effectiveness in handling complex, multi-variable scheduling constraints. The GA is particularly suitable for exploring a wide range of scheduling combinations, allowing it to identify stable solutions that consistently minimize wait times across various scheduling scenarios. Conversely, BO is well-suited to efficiently allocate resources by leveraging probabilistic models to refine scheduling parameters iteratively. Both methods aim to reduce patient wait times by minimizing the value derived from Eq. (1). An optimal solution is defined mathematically as the schedule configuration that minimizes the fitness function described in Eq. (1) (global average wait time), subject to hard constraints on room eligibility (e.g., fixed-beam patients must use the fixed-beam room). Convergence is considered achieved when the average wait time does not improve for a set number of iterations, or when the maximum iteration count is reached. All simulations and optimizations were performed on a standard clinical workstation (Intel Core i7 processor, 16 GB RAM) to demonstrate feasibility for use in most radiation therapy centers.

Genetic algorithm

The genetic algorithm (GA) is a population-based optimization technique inspired by natural selection [18]. It was employed to enhance patient scheduling and treatment room allocation iteratively. The algorithm begins by initializing the population with random parameter values within predefined ranges for the number of patients, treatment rooms, probability of needing a fixed-beam room, and arrival intervals. Each candidate's quality is evaluated using a fitness function based on the average wait time, calculated through simulation for each patient order and room assignment.

The evolutionary process in the GA involves selection, crossover, and mutation [20], which work together to refine the population. Selection is performed using a tournament selection method [21], where subsets of the population are compared, and the highest-performing individuals are selected. Crossover is then applied to combine characteristics from parent solutions, allowing exploration of new possibilities. Mutation randomly adjusts parameters and shuffles the patient order in a small number of solutions, promoting diversity and preventing premature convergence on suboptimal solutions. Over the course of generations, the population evolves through repeated application of selection, crossover, and mutation, continually improving the solution quality. The genetic algorithm treats the patient schedule as a permutation chromosome. To ensure valid schedules during evolution, we utilized Order Crossover (OX1), which preserves the relative ordering of patient sub-sequences from one parent while filling remaining slots from the other, thus preventing patient duplication or omission. Swap Mutation was employed with a probability of 0.05 to introduce diversity by randomly exchanging two patients in the sequence.

The fitness function, which takes the negative of the average wait time, is used to evaluate each solution, ensuring that lower wait times result in higher fitness scores. The algorithm was configured with a computational budget of 30,000 function evaluations. Convergence was defined as 1500 evaluations of stagnation. This optimization minimizes patient wait times by improving the allocation of treatment resources and enhancing scheduling efficiency.

Algorithm 1. Genetic algorithm (GA) optimization for patient schedule

Inputs:
P_{range}: Range of number of patients.
R_{range}: Range of number of treatment rooms.
P_{fixed}: Probability range for needing a fixed-beam room.
$A_{interval}$: Range for mean arrival interval of patients.
Outputs:
W_{init}: Initial average wait time.
P_{best}: Best set of input parameters.
O_{best}: Optimized patient order.
W_{opt}: Average optimized wait time.
Determine initial parameters (P, R, P_{fixed}, $A_{interval}$) within specified ranges.
Execute simulation with (P, R, P_{fixed}, $A_{interval}$) to calculate W_{init}.
Calculate fitness score based on W_{init}.
Designate the initial schedule as the “Parent.”
for *generation* = 1 to 100 **do**
Modify R and P_{fixed} within their ranges.
Shuffle $O_{current}$ to generate O_{new}.
Combine $O_{current}$ from two parent schedules to create $O_{offspring}$.
Calculate fitness of $O_{current}$ based on new wait time.
Choose the best schedules (O_{best}, P_{best}) from the current generation.
end for
Return W_{init}, P_{best}, O_{best}, and W_{opt}.

Bayesian optimization

Bayesian optimization (BO) [21] is an effective technique known for its ability to efficiently explore high-dimensional search spaces and handle computationally expensive objective functions. In this study, we employed BO to refine the same objective function as in the GA, aiming to minimize the average patient wait time.

BO works by iteratively exploring the parameter space through the construction of a probabilistic model of the objective function, using a Gaussian Process [22] with Matern kernel [23]. This model is continuously updated as new evaluations are performed, allowing the optimizer to better predict the performance of untested parameter combinations. The optimization process is guided by an acquisition function [24], which manages the balance between exploring uncertain regions of the parameter space and focusing on areas where the model predicts optimal results. The acquisition function selects the next set of parameters to evaluate based on both the model's uncertainty and the potential for improvement.

The BO process begins with 20 initial evaluations using randomly selected parameter combinations, including the number of patients, treatment rooms, the probability of requiring a fixed-beam treatment room, and the mean patient arrival interval. These initial evaluations provide a baseline for model predictions, contributing to establishing a reference point for the Gaussian Process to predict performance across untested parameter spaces. In each subsequent iteration, the BO selects a new set of parameters informed by feedback from prior evaluations, which dynamically updates the Gaussian Process model. This model refinement enhances the optimizer's capacity to balance exploration of uncertain areas and exploitation of promising parameters. The process for BO in the context of the patient schedule is outlined in Algorithm 2.

Algorithm 2. Bayesian optimization (BO) for patient schedule

Inputs:
P_{range}: Range of number of patients.
R_{range}: Range of number of treatment rooms.
P_{fixed}: Probability range for needing a fixed-beam room.
$A_{interval}$: Range for mean arrival interval of patients.
Outputs:
W_{init}: Initial average wait time.
P_{best}: Best set of input parameters.
O_{best}: Optimized patient order.
W_{opt}: Average optimized wait time.
Randomly select initial parameters (P, R, P_{fixed}, $A_{interval}$) within specified ranges.
Compute 20 simulations with (P, R, P_{fixed}, $A_{interval}$) to calculate W_{init}.
Calculate fitness score based on W_{init}.
Use initial data to build a probabilistic model of the objective function.
for *iteration* = 1 to 2 0 **do**
Perform initial evaluations using randomly select initial parameters (P, R, P_{fixed}, $A_{interval}$) to establish a reference point for Gaussian Process model.
for *iteration* = 1 to 80 **do**
Use an acquisition function to select next parameters (P, R, P_{fixed}, $A_{interval}$).
Generate patient order O.
Run simulation with (P, R, P_{fixed}, $A_{interval}$, O) to calculate wait time W.
Update the Gaussian Process model with new data.
end for
Find P_{best} and O_{best} that minimizes W.
Calculate W_{opt} as the optimized average wait time.
Return W_{init}, P_{best}, O_{best}, and W_{opt}.

The iterative process continues for a maximum of 80 iterations, for a total of 100 initial evaluations and subsequent adaptations, the same number used for the GA.

Results

Table 2 provides a comparative performance analysis of the GA and BO in minimizing average patient wait times over 100 simulation repetitions. The baseline average wait time was 0.45 min, with a 95% confidence interval of 0.38–0.52 min.

Following 100 generations, the GA achieved an average wait time of approximately 0.19 min (9.6 s), highlighting its optimization capacity. Each GA completed within an average execution time of 17.14 s. Fig. 1 shows a patient schedule before and after optimization with GA and BO algorithms.

In contrast, the BO approach, utilizing a Matern kernel within the Gaussian Process, was executed with a total of 100 iterations. BO slightly outperformed the GA, minimizing patient wait times to an average of approximately 0.11 min. Notably, BO demonstrated superior computational efficiency, requiring an average execution time of only 0.36 s per simulation, approximately 48 times faster than the GA.

Table 2
Performance comparison of genetic algorithm (GA) and Bayesian optimization (BO) in minimizing patient wait times across 100 repeated simulations.

Optimization Method	Average Patient Wait Time (100 simulations) (95% CI, min)	Average Execution Time (100 simulations) (95% CI, sec)
Baseline	0.45 (0.38–0.52)	N/A
Genetic Algorithm (GA)	0.19 (0.14–0.25)	17.14 (15.47–18.80)
Bayesian Optimization (BO)	0.11 (0.08–0.14)	0.36 (0.33–0.39)

Abbreviations: 95% CI, 95% confidence interval; N/A, Not applicable; min, Minutes; sec, Seconds.

Sensitivity analysis

To evaluate the robustness of our optimization approaches, we conducted sensitivity analysis across 50 diverse test scenarios. Both algorithms demonstrated consistent performance improvements across all tests, with the GA achieving wait time reductions averaging 58% (95% CI: 0.14–0.25 min) and BO achieving reductions averaging 76% (95% CI: 0.08–0.14 min).

The narrow confidence intervals for both algorithms indicate stable performance across varying test conditions. The GA demonstrated superior optimization in 39% of test cases, while BO achieved best results in 46% of cases, with 24% resulting in effective ties, suggesting complementary strengths between the two approaches.

Statistical significance testing confirmed that BO had significantly faster execution times than the GA (mean difference of 16.78 s, $p < 0.001$), representing a 48-fold improvement in computational efficiency.

Outlier and frequency analysis

To assess clinical impact beyond average performance, we utilized a two-tiered statistical approach. First, a detailed outlier analysis of 100 simulations quantified the reduction in extreme delays. The baseline schedule resulted in an average maximum wait time of 2.30 min and a 95th percentile wait time of 1.68 min. Bayesian Optimization (BO) significantly truncated these extremes, reducing the average maximum wait time to 1.69 min ($p < 0.001$) and the 95th percentile to 1.12 min ($p < 0.001$).

To assess robustness, we expanded this to a frequency analysis of 1000 simulations. Statistical significance for each simulation was determined using a one-tailed Mann-Whitney *U* test to confirm a reduction in wait times. The results showed that BO achieved statistically significant reductions ($p < 0.05$) in mean wait times in 32.3% of all randomized scenarios and significant reductions in maximum wait times in 28.9% of scenarios. This indicates that the algorithm is highly effective at identifying and correcting schedules with significant inefficiencies, while maintaining stability in schedules that are already near-optimal.

Discussion

This study aims to improve patient scheduling in heavy ion therapy facilities by developing optimization algorithms that address the unique scheduling challenges of managing multiple treatment rooms competing for a single beam source. By integrating simulation with optimization techniques, we enhance patient scheduling and resource allocation while evaluating the efficacy of two optimization algorithms. Our simulation–optimization model incorporated realistic variables such as patient arrival rates, tumor types, treatment times, and the probability of requiring a fixed-beam room. By applying the GA and BO, we significantly reduced the average patient wait times compared to the baseline, demonstrating the effectiveness of these algorithms in improving scheduling efficiency.

The baseline average wait time of 0.45 min reflects the inherent inefficiencies in current scheduling practices within heavy ion therapy facilities. These prolonged wait times can lead to patient discomfort, decreased satisfaction, and suboptimal utilization of expensive resources. Both optimization methods substantially reduced the average wait times, down to 0.19 min with GA and 0.11 min with BO, highlighting their potential to enhance patient experience and operational efficiency. Although an absolute saving of 20–30 s per patient may appear clinically insignificant, this overlooks two critical factors: patient experience and cumulative facility throughput. First, this wait time is immobilized time on the treatment couch. For patients in masks or other complex immobilization devices, reducing this period of discomfort and anxiety, even by seconds, is a direct and meaningful improvement to the quality of patient care. Second, and most critically, these small individual savings accumulate significantly over a full operational day. It is important to distinguish between total clinic wait time (appointment to room entry) and in-room beam wait time. While the former impacts general satisfaction, the latter represents a critical quality-of-care metric for heavy ion therapy. Delays while immobilized in a thermoplastic mask or vacuum cushion can cause significant anxiety and physical discomfort compared to delays in a waiting room. Therefore, our objective function specifically targets the minimization of this immobilized interval.

The GA showed strong optimization capabilities, winning 39% of head-to-head comparisons. However, BO demonstrated notable efficiency, achieving comparable or slightly better average wait times (0.11 min versus 0.19 min) while requiring only 0.36 s of execution time compared to 17.14 s for the GA. This significant computational advantage (approximately 48 times faster) makes BO particularly attractive for real-time scheduling applications where rapid decisions are critical. The optimization was performed on a standard clinical workstation. The low computational cost, particularly of BO (0.36 s), allows these algorithms to be deployed without the need for high-performance computing clusters. The performance characteristics suggest different use cases: GA might be preferred when computation time is less constrained and finding the absolute optimal solution is paramount, while BO excels in time-sensitive scenarios requiring quick, high-quality solutions. Both approaches effectively minimized patient wait times from baseline; however, BO demonstrated faster convergence and a more consistent reduction in wait times compared to GA. This suggests complementary use cases: BO's 48-fold speed advantage makes it ideal for real-time, on-the-fly schedule adjustments during the clinic day. In contrast, the GA's performance in a significant minority of cases suggests its utility for offline, weekly schedule planning, where more computation time can be dedicated to finding a globally robust schedule.

Our findings substantially exceed previous research in the field of healthcare scheduling optimization. It is important to note that our study is a simulation-based proof-of-concept and direct comparison is therefore limited by differences in facility configuration, patient populations, and the presence of operational complexities not captured in simulation. Wang et al [16] achieved about a 35% reduction in wait times using their RGCC method in a proton therapy facility, while our approaches demonstrated improvements of 29.09% (GA) and 76.95% (BO). This could be due to our consideration of additional parameters and the inherent advantages of BO in exploring complex solution spaces. It is also likely that different constraints were placed on their simulation compared to the one presented in this work. The present simulation was for a mixed carbon ion and proton beam facility, with additional patients that would only be treated using carbon ion therapy. While the application of GA and BO to scheduling is an established field, the novelty of this work lies in its application to a realistic simulation model incorporating the unique, complex constraints of a multi-gantry CIRT facility (e.g., shared synchrotron, mixed fixed-beam/gantry rooms, and mid-treatment setups), which has not been well-studied. Direct collaboration with treatment delivery professionals is an essential next step to refine the model's assumptions and ensure its practical utility.

Implementation of our optimization approach in real clinical settings would require integration with existing hospital information systems and treatment planning workflows. Key considerations would include staff training, development of user-friendly interfaces, and establishment of protocols for handling schedule adjustments during unforeseen circumstances such as patient delays or equipment malfunctions. The exceptionally rapid execution time of the BO algorithm (0.36 s) makes it particularly suitable for real-time schedule updates, potentially allowing facilities to dynamically adjust to changing conditions throughout the treatment day with minimal computational overhead.

Despite the promising results, our study has several limitations. The primary limitation is reliance on simulated data rather than empirical operational data. While parameters are derived from published literature, the simulation necessarily simplifies real-world complexity. The performance of the optimization algorithms might differ when applied to real-world data with unforeseen complexities. Future studies should aim to validate and refine the model using empirical data from operating carbon ion therapy facilities to enhance its accuracy and reliability.

Additionally, the focus of this study was on minimizing patient wait times within the treatment room, without explicitly considering other critical factors such as staff scheduling, equipment maintenance, emergency cases, and patient-specific needs that could impact the overall scheduling. Incorporating these elements into the optimization framework could provide a more comprehensive approach to improving operational efficiency and patient care.

As a proof of concept, this simulation assumes an uninterrupted operational flow to isolate the effects of scheduling logic. Consequently, common clinical disruptions were excluded from the current model, such as beam interlocks, synchrotron downtime, anesthesia induction delays, and patient transport logistics. Real-world implementation would likely see higher absolute wait times due to these factors, though the relative benefit of optimization in resolving resource conflicts remains valid.

Furthermore, while BO outperformed the GA in terms of average wait time reduction and computational speed, the choice of optimization algorithm may depend on specific operational constraints and objectives of a facility. For instance, facilities with ample computational resources might benefit from Bayesian exploration of the parameter space, potentially leading to solutions that are more robust under varying conditions. It should be noted that intensive optimized schedules must not sacrifice patient safety or quality for efficiency.

Our approach produces optimized static schedules before the treatment day begins. Future research directions include exploring machine learning techniques to predict scheduling bottlenecks and dynamically adjust parameters in real-time. Additionally, extending the optimization objectives to include multiple criteria such as machine downtime and staff workload balance could provide a more holistic improvement to the scheduling system.

Conclusions

In conclusion, the implementation of optimization algorithms, particularly the BO algorithm, in the scheduling of multi-room heavy ion therapy facilities shows significant potential in reducing immobilized patient wait times. The clinical significance extends beyond the absolute time savings to encompass improved patient comfort, enhanced facility throughput, and better utilization of resources. By addressing the unique challenges of heavy ion therapy scheduling, our study contributes valuable insights that can aid in the development of more efficient and patient-centered treatment facilities. As CIRT becomes more widely adopted, the application of such optimization strategies will be essential in maximizing the benefits of this advanced treatment modality for both patients and healthcare providers.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Research support was provided by the University of Chicago Medicine Comprehensive Cancer Center Trainee Associate Member Professional Development Award.

References

[1] Wilson RR. Radiological use of fast protons. Radiology 1946;47(5):487–91. https://doi.org/10.1148/47.5.487.
[2] Lawrence JH. Proton irradiation of the pituitary. Cancer 1957;10(4):795–8. https://doi.org/10.1002/1097-0142(195707/08)10:4.
[3] Particle Therapy Co-Operative Group. Particle Therapy Facilities in Clinical Operation. Published 2019. Accessed April 21, 2025. https://www.ptcog.site/index.php/facilities-in-operation.
[4] Collings EW, Lu L, Gupta N, Sumption MD. Accelerators, gantries, magnets and imaging systems for particle beam therapy: recent status and prospects for improvement. Front Oncol 2022;11:737837. https://doi.org/10.3389/fonc.2021.737837.
[5] Beltran C, Amos RA, Rong Y. We are ready for clinical implementation of carbon ion radiotherapy in the United States. J Appl Clin Med Phys 2020;21(12):6. https://doi.org/10.1002/acm2.13133.
[6] Malouff TD, Mahajan A, Krishnan S, Beltran C, Seneviratne DS, Trifiletti DM. Carbon ion therapy: a modern review of an emerging technology. Front Oncol 2020;10:82. https://doi.org/10.3389/fonc.2020.00082.
[7] Bottura L, Felcini E, Ferrero V, Fiorina E, Monaco V, Pennazio F, et al. Delivery, beam and range monitoring in particle therapy in a highly innovative integrated design. Front Phys 2020;8:566679. https://doi.org/10.3389/fphy.2020.566679.
[8] Chinniah S, Deisher AJ, Herman MG, Johnson JE, Mahajan A, Foote RL. Rotating gantries provide individualized beam arrangements for charged particle therapy. Cancers 2023;15(7):2044. https://doi.org/10.3390/cancers15072044.
[9] Yagi M, Beltran CJ, Shimizu S, Hamatani N, Tsubouchi T, Takashina M, et al. Carbon ion therapy for laterally located tumors require multiple fixed ports or a rotating gantry. Med Dosim 2024;49(4):279–90. https://doi.org/10.1016/j.meddos.2024.02.003.
[10] Burns M, Campbell R, French S, Dhillon HM, Butow PN, Pritchard A, et al. Trajectory of anxiety related to radiation therapy mask immobilization and treatment delivery in head and neck cancer and radiation therapists' ability to detect this anxiety. Adv Radiat Oncol 2022;7(5):100967. https://doi.org/10.1016/j.adro.2022.100967.
[11] Forbes E, Clover K, Baker AL, Britton B, Carlson M, McCarter K. 'Having the mask on didn't worry me until they clamped my head down so I wouldn't move': a qualitative study exploring anxiety in patients with head and neck cancer during radiation therapy. J Med Radiat Sci 2023;70(3):283–91. https://doi.org/10.1002/jmrs.658.
[12] Nixon JL, Brown B, Pigott AE, Turner J. A prospective examination of mask anxiety during radiotherapy for head and neck cancer and patient perceptions of management strategies. J Med Radiat Sci 2019;66(3):184–90. https://doi.org/10.1002/jmrs.346.
[13] Huang YL, Deisher AJ, Herman MG, Kruse JJ, Mahajan A. Reduce patient treatment wait time in a proton beam facility–a gatekeeper approach. J Med Syst 2021;45(8):80. https://doi.org/10.1007/s10916-021-01756-1.
[14] Wang F, Huang Y-L, Ju F. Simulation optimization gantry call-back control method for proton therapy systems. IEEE Trans Autom Sci Eng 2023;20(3):1565–76. https://doi.org/10.1109/TASE.2022.3198106.
[15] Wang F, Huang Y-L, Ju F. (2021). Rollout-based Gantry Call-back Control for Proton Therapy Systems. In 2021 IEEE 17th International Conference on Automation Science and Engineering (CASE) (pp. 962–967). IEEE. https://doi.org/10.1109/CASE49439.2021.9551416.
[16] Zinoviev D. (2024). Discrete event simulation: it's easy with SimPy!. *arXiv*. https://doi.org/10.48550/arXiv.2405.01562.
[17] Suzuki K, Palmer MB, Sahoo N, Zhang X, Poenisch F, Mackin DS, et al. Quantitative analysis of treatment process time and throughput capacity for spot scanning proton therapy. Med Phys 2016;43:3975–86. https://doi.org/10.1118/1.4952731.
[18] Alhijawi B, Awajan A. Genetic algorithms: theory, genetic operators, solutions, and applications. Evol Intel 2024;17(3):1245–56. https://doi.org/10.1007/s12065-023-00822-6.
[19] Frazier PI. (2018). A tutorial on Bayesian optimization. arXiv preprint arXiv:1807.02811. https://doi.org/10.48550/arXiv.1807.02811.
[20] Hassanat A, Almohammadi K, Alkafaween E, Abunawas E, Hammouri A, Prasath VS. Choosing mutation and crossover ratios for genetic algorithms—a review with a new dynamic approach. Information 2019;10(12):390. https://doi.org/10.3390/info10120390.
[21] Pandey HM, Shukla A, Chaudhary A, Mehrotra D. (2016). Evaluation of genetic algorithm's selection methods. In Information Systems Design and Intelligent Applications: Proceedings of Third International Conference INDIA 2016, Volume 2 (pp. 731–738). Springer India. https://doi.org/10.1007/978-81-322-2752-6_72.
[22] Lizotte DJ, Wang T, Bowling MH, Schuurmans D. (2007). Automatic gait optimization with gaussian process regression. In Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI) (Vol. 7, pp. 944–949).
[23] Minasny B, McBratney AB. The Matérn function as a general model for soil variograms. Geoderma 2005;128(3–4):192–207. https://doi.org/10.1016/j.geoderma.2005.04.003.
[24] Jones DR, Schonlau M, Welch WJ. Efficient global optimization of expensive black-box functions. J Glob Optim 1998;13:455–92. https://doi.org/10.1023/A:1008306431147.