

Mixed integer programming with dose-volume constraints in intensity-modulated proton therapy

Pengfei Zhang¹ | Neng Fan² | Jie Shan³ | Steven E. Schild¹ | Martin Bues¹ | Wei Liu¹

¹Department of Radiation Oncology, Mayo Clinic, Scottsdale, AZ, USA

²Department of Systems & Industrial Engineering, University of Arizona, Tucson, AZ, USA

³Department of Biomedical Informatics, Arizona State University, Phoenix, AZ, USA

Author to whom correspondence should be addressed. Wei Liu

E-mail: liu.wei@mayo.edu; Telephone: (480) 342-4215

Funding Information

National Cancer Institute (NCI), Grant/Award Number: K25CA168984

Abstract

Background: In treatment planning for intensity-modulated proton therapy (IMPT), we aim to deliver the prescribed dose to the target yet minimize the dose to adjacent healthy tissue. Mixed-integer programming (MIP) has been applied in radiation therapy to generate treatment plans. However, MIP has not been used effectively for IMPT treatment planning with dose-volume constraints. In this study, we incorporated dose-volume constraints in an MIP model to generate treatment plans for IMPT.

Methods: We created a new MIP model for IMPT with dose volume constraints. Two groups of IMPT treatment plans were generated for each of three patients by using MIP models for a total of six plans: one plan was derived with the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) method while the other plan was derived with our MIP model with dose-volume constraints. We then compared these two plans by dose-volume histogram (DVH) indices to evaluate the performance of the new MIP model with dose-volume constraints. In addition, we developed a model to more efficiently find the best balance between tumor coverage and normal tissue protection.

Results: The MIP model with dose-volume constraints generates IMPT treatment plans with comparable target dose coverage, target dose homogeneity, and the maximum dose to organs at risk (OARs) compared to treatment plans from the conventional quadratic programming method without any tedious trial-and-error process. Some notable reduction in the mean doses of OARs is observed.

Conclusions: The treatment plans from our MIP model with dose-volume constraints can meet all dose-volume constraints for OARs and targets without any tedious trial-and-error process. This model has the potential to automatically generate IMPT plans with consistent plan quality among different treatment planners and across institutions and better protection for important parallel OARs in an effective way.

PACS

87.55.de, 87.55.kd, 81.40.Wx

KEY WORDS

dose-volume constraints, intensity-modulated proton therapy, mixed-integer programming

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2017 The Authors. *Journal of Applied Clinical Medical Physics* published by Wiley Periodicals, Inc. on behalf of American Association of Physicists in Medicine.

1 | INTRODUCTION

Proton therapy, which has been used to treat cancer since the 1950s, offers many clinical advantages compared with the conventional radiation therapy, such as x rays or electron beams. Compared with intensity-modulated radiation therapy (IMRT), intensity-modulated proton therapy (IMPT) can deliver highly conformal dose distributions to the target yet spare organs at risk (OARs). Therefore, IMPT can greatly reduce the risk of OAR damage and increase the chance of local control of cancer.^{1,2}

In treatment planning for IMPT, we aim to deliver the prescribed dose to the target yet minimize the dose to adjacent OARs. Dose-volume constraints are used in most modern, commercial treatment planning systems, such as Raystation (RaySearch Laboratories), Eclipse (Varian Medical Systems), and Pinnacle (Philips Radiation Oncology Systems, Philips Healthcare), to generate clinically acceptable treatment plans. These treatment planning systems all use quadratic programming models for optimization and many reports have described their use in radiotherapy. Wu and Mohan³ implemented soft dose-volume constraints by quadratic programming, with the dose-volume constraints in the objective function. Falkinger et al.⁴ proposed a prioritized optimization algorithm for IMPT planning. Gradient-based methods³ and a genetic algorithm⁵ have also been used to generate treatment plans. Xing et al.⁶ reported a method for estimating the parameters for the nonlinear objective function and the corresponding algorithm used to determine those parameters. Chen et al.⁷ developed a database-generation procedure for IMPT treatment planning.

These methods all use soft constraints and require input of initial optimization parameters, for example, penalty weights, to achieve the best balance between tumor coverage and normal tissue protection. As a result of the soft constraints applied in those methods, the desired dose volume constraints may not be rigidly satisfied in the generated results. In addition, the parameters they used are determined based on the planner's experience. Therefore, many time-consuming trial-and-error iterations are needed to get the optimal initial optimization parameters. Ferris et al.⁸ showed that trial-and-error methods for IMRT are impractical because of the complexity of treatment planning and that use of mixed integer programming (MIP) is a more efficient method to ensure that the treatment plan meets the clinical requirements. In addition, hard constraints used in linear programming minimize the variation in plan quality among different treatment planners and across institutions, which is important for multiple institution clinical trials.

Linear programming can model clinic requirement in a mathematical model by linear relationships. Linear programming is a special case of mathematical programming, which is to optimize a linear objective function, subject to linear equality and linear inequality constraints. Additionally, we need to introduce integer variables to deal with dose volume constraints since patients' geometry is voxelized. Therefore, mixed integer programming is a very good way for IMPT problem. Compared to quadratic programming, we can use linear programming to explicitly implement hard constraints for IMPT

with the cost of complexity of models and more extensive computation time. Many investigators have studied the use of MIP in radiotherapy. For example, Cao et al.⁹ used MIP for optimizing the beam angle to deal with the uncertainties in various scenarios. To improve the plan quality in IMPT, Zaghian et al.¹⁰ proposed an iterative approach that could satisfy the dose-volume constraints in MIP. Tuncel et al.¹¹ derived strong inequalities for MIP in IMRT. Rocha et al.¹² proposed a programming approach for IMRT using binary integers and obtained a plan with improved quality. Romeijn et al.¹³ proposed an MIP model for IMRT, in which the dose-volume constraints are explicitly enforced. However, scant research exists regarding how to apply MIP to IMPT treatment planning with dose-volume constraints (except Zaghian et al.¹⁰).

In this study, we propose a new method to apply the MIP model with dose-volume constraints to the IMPT treatment planning. In the MIP model, we can specify hard dose volume constraints on tumors and OARs. The generated treatment plan can satisfy all the dose volume constraints for tumors and OARs if any feasible solution exists.

In our MIP model we will specify the dose volume constraints for tumors and OARs in the optimization. If the tumor constraints (tumor coverage and homogeneity) are too stringent, there may not be a feasible solution. On the other hand, if the tumor constraints are too loose, the resultant plan is not optimal in tumor dose distribution even though it meets all the requirements for all normal tissue protection. Therefore, given the same normal tissue constraints, it is important to find the best tumor dose volume constraint parameters in MIP. Again one could use trial-and-error process to get as good parameters as possible. However, it is time consuming and the best result is not guaranteed. In this work, we have further developed a new method to automatically find the best parameters of tumor dose volume constraints in MIP to solve this problem.

2 | MODEL AND FORMULATION

2.A | Patient data and beam configurations

We retrospectively generated IMPT plans for three patients: one pediatric patient with prostate rhabdomyosarcoma (prescription dose, 45 Gy) and two patients with head and neck (H&N) cancer (prescription dose, 60 Gy or 68 Gy). These three cases are typical clinic cases, and in practice it is hard to generate good plans to meet all the dose volume constraints for these three cases due to the close proximity of critical organs to tumors. We have chosen them to demonstrate the effectiveness of our model. The corresponding dose-volume constraints used in our clinic are shown in Table 1. The dose covering a percentage of the structure's volume ($D_{\%}$) derived from the structure's dose-volume histogram (DVH) was compared. The planning target volume (PTV) was formed by uniform expansion of the clinical target volume (CTV) by 5 mm for the prostate rhabdomyosarcoma case and by 3 mm for the H&N cases. The PTV $D_{95\%}$ and $D_{5\%}-D_{95\%}$ were used to assess tumor dose coverage and homogeneity, respectively. Three beams were used for all three patients. The number of beamlets in each structure is also included

TABLE 1 Dose-volume constraints.

Structure	Dose volume constraints (Gy[RBE])	Achieved values from MIP (Gy[RBE])	Achieved values from L-BFGS (Gy[RBE])
Prostate (with 3837 beamlets)			
Planning target volume $D_{95\%}$	>44	44	44
Planning target volume $D_{5\%}-D_{95\%}$	<2	1.6	2
Bladder $D_{25\%}$	<45	44.2	44.5
Femoral heads $D_{10\%}$	<28	12.4	14.1
Rectum $D_{25\%}$	<45	45	44.8
H&N, case 1 (with 7773 beamlets)			
Planning target volume $D_{98\%}$	>59	60	59.3
Planning target volume $D_{2\%}-D_{98\%}$	<2	0	1.3
Whole brain $D_{1\%}$	<62	60	59.5
Brain stem $D_{1\%}$	<55	55	53
Optic chiasm $D_{1\%}$	<55	55	56.1
H&N, case 2 (with 9905 beamlets)			
Planning target volume $D_{98\%}$	>67	68	67.6
Planning target volume $D_{2\%}-D_{98\%}$	<2	0	0.8
Brain stem $D_{1\%}$	<55	25.2	24.3
Mucosa avoid $D_{1\%}$	<30	15.4	16.3
Spinal cord $D_{1\%}$	<45	20.2	21.9

in Table 1. A resolution of 5 mm was used in the dose calculation and optimization.

For each patient, we used two optimization methods to generate IMPT plans with identical dosimetric goals: the MIP model and quadratic programming method. The optimization software, IBM CPLEX (version 12.5)¹³ was used to solve the MIP problems. We also generated IMPT plans for the same patients using a quadratic programming model solved by the Limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method. L-BFGS is a limited-memory quasi-Newton code for unconstrained optimization. L-BFGS is similar to the classical gradient descent method, but it uses some approximation to minimize the memory overhead. Therefore it is particularly well suited for optimization problems with a large number of variables. IMPT treatment planning involves a huge number of variables (usually as large as 50,000), thus it is a good application for L-BFGS. Currently the new version of the commercial treatment planning system, EclipseTM, has an option to use L-BFGS as an optimizer. Thus we believe that it is a good idea to benchmark our new method with the L-BFGS method. Please note that we used the same beamlets (influence matrix) for both quadratic programming and MIP for fair comparison.

2.B | Implementation of dose-volume constraints

For each voxel i within each structure $l \in \{OAR, T\}$ (tumors are denoted as T), D_0^l and D_i^l denote the prescribed dose value and the plan dose, respectively, and P^l is the percentage of the structure volume covered by a certain prescribed dose D_0^l . In addition, N^l denotes the total number of voxels of the structure l , and D_0^{TL} (upper) and D_0^{TR} (lower) are the prescribed doses covering the corresponding P^{TL} (upper) and P^{TR} (lower) percentage of the volume of the target, respectively. For example, by specifying the P^{TR} to be 0.95, the D_0^{TR} becomes $D_{95\%}$. The binary decision variables $y_i \in \{0, 1\}$ denote whether the dose at the voxel i is less than the prescribed dose value ($y_i = 0$) or not ($y_i = 1$), and $y = (y_1, y_2, \dots, y_n)$ denotes the binary decision variables for all voxels ($i = 1, 2, \dots, n$).

The dose-volume constraints are specified for normal organs or for hot spots and cold spots in tumors. The dose-volume constraints limit the percentage of the structure volumes with doses exceeding the prescribed dose value to be less than or equal to the specified value of P^L or P^{TR} . In addition, to ensure the specified dose-volume constraints are met for cold spots in tumors, we limit the percentage of the tumor volumes that receive lower doses than the prescribed dose value to be less than or equal to the specified value of P^{TL} . The MIP formulation follows (called **Model 1** hereafter),

$$\min \sum_{i \in T} |D_i^T - D_0^T| \quad (1a)$$

$$\text{s.t. } D_i^l \leq D_0^l + M y_i^l, \forall i \in l, \forall l \in OAR, \quad (1b)$$

$$D_i^l \geq D_0^l y_i^l, \forall i \in l, \forall l \in OAR, \quad (1c)$$

$$\sum_{i \in T} y_i^T / N^T \leq P^L, \forall l \in OAR, \quad (1d)$$

$$D_i^T \leq D_0^{TR} + M y_i^{TR}, \forall i \in T, \quad (1e)$$

$$D_i^T \geq D_0^{TR} y_i^{TR}, \forall i \in T, \quad (1f)$$

$$\sum_{i \in T} y_i^{TR} / N^T \leq P^{TR}, \quad (1g)$$

$$D_i^T \leq D_0^{TL} + M(1 - y_i^{TL}), \forall i \in T, \quad (1h)$$

$$D_i^T \geq D_0^{TL}(1 - y_i^{TL}), \forall i \in T, \quad (1i)$$

$$\sum_{i \in T} y_i^{TL} / N^T \leq P^{TL}, \quad (1j)$$

$$D_i^T = \sum_j w_{ij}^T x_j, \forall i \in T \quad (1k)$$

$$D_i^l = \sum_j w_{ij}^l x_j, \forall i \in l, \forall l \in OAR \quad (1l)$$

$$y_i^l \in \{0, 1\}, \forall i \in l, \forall l \in OAR, \quad (1m)$$

$$y_i^{TL}, y_i^{TR} \in \{0, 1\}, \forall i \in T \quad (1n)$$

$$x_j \geq 0, \forall j \in J \quad (1o)$$

where M is a sufficiently large positive number. The role of M is to make the inequality hold when the term associated with M takes a value of 1 (i.e., $y_i = 1$). We can choose M to be the upper bound of all possible D_0^l for OARs and D_0^{TR} for targets. For the cases included in this research, it is set to be 100 Gy[RBE]. It is worth noting that it is better not to set M too large, because as M becomes larger, all the constraints, which contain M , become less contingent, and thus the computation time becomes longer. The objective (1a) is to

minimize the total dose deviation on tumors. Instead of the quadratic objective function, we used the absolute value to compute total dose deviation. Constraints (1b) and (1c) together are to decide whether the dose at a voxel i of a normal organ l exceeds the prescribed dose D_0 or not. If the dose at a voxel i of a normal organ l does not exceed the prescribed dose D_0 , then the corresponding binary variable y_i is 0. Otherwise, the corresponding binary variable y_i is 1. This is facilitated by a sufficiently large positive number M . The role of M is to ensure that the constraints hold when y variable is 1. Constraints (1d) are the dose-volume constraints for OARs with the help of the binary variable y_i , determined by (1b) and (1c). We limit the percentage of the organ l volumes with doses exceeding D_0^l to be less than or equal to P^l . Similarly, constraints (1e) to (1g) are the dose-volume constraints for tumor hot spots. We limit the percentage of the tumor volumes with doses exceeding D_0^{TR} to be less than or equal to P^{TR} . Constraints (1h) to (1j) are for the dose-volume constraints for tumor cold spots. Equations (1k) and (1l) represent the total dose to each voxel from all beamlets, and x_j are continuous variables representing the j th beamlet intensity (fluence), while the influence matrix, w_{ij} , is the dose contribution of the beamlet j of unit intensity at voxel i . The influence matrix is precalculated using an in-house developed IMPT dose calculation engine.^{14,15} Constraints (1m) and (1n) are the constraints for the binary decision variables.

Note that in the objective function, we only include terms for tumors but no terms for OARs, this is because all desired constraints for OARs are included in the constraints in the model (unlike quadratic programming, the dose volume constraints are usually included in the objective function as “soft constraints”). Note that in our model, we have products of continuous variables and binary variables. This kind of nonlinear terms can be linearized by introducing one additional continuous variable and additional constraints. For example, we can linearize $D_0^l y_i^l$ by introducing variable z_i^l , constant M (large enough) and following constraints:

$$z_i^l \geq D_0^l y_i^l, z_i^l \leq D_0^l, z_i^l \geq 0, z_i^l \leq M y_i^l$$

We have modeled all requirements as hard constraints. In practice, this can ensure all the dose volume constraints. Usually, target coverage may be satisfied to satisfy constraints for some critical organs such as spinal cord and brain stem. In such situation, we can always change the parameters in the constraints to achieve the best balance with the “trial-and-error” procedure.

2.C | Minimizing the trial-and-error procedure

There is a drawback to our MIP model with dose volume constraints. After we set the dose-volume constraints for structures, the original optimization problem (Model 1) might not have feasible solutions. We have to try different values for $D_{5\%}$ and $D_{95\%}$ to find the best balance between target coverage and protection of OARs, which can ensure good homogeneity. If the values are too stringent, Model 1 becomes infeasible; if the values are too loose, the resulting plan is not optimal for the tumor dose distribution. Therefore, it is

important to find ways to obtain the best values of D_0^{TR} and D_0^{TL} so that we obtain the optimal balance.

One potential way to acquire optimal parameters is to use the trial-and-error procedure by solving Model 1 with different values for D_0^{TR} and D_0^{TL} ; however, this procedure is tedious and optimal results are not guaranteed. In this section, we propose the following method to find the best parameters for tumors given the same dose volume constraints for normal tissues. In the following explanation, we are going to take the prostate case as an example. If $P^{TR} = P^{TL} = 0.05$, the D_0^{TR} becomes $D_{95\%}$, and the D_0^{TL} becomes $D_{5\%}$. We propose a new way to obtain the best parameters more efficiently (called **Model 2** hereafter) to minimize the trial-and-error process as follows,

$$\min d_i^T - d_i^T \quad (2a)$$

$$\text{s.t. } D_i^l \geq D_0^l + M y_i^l, \forall i \in I, \forall l \in \text{OAR}, \quad (2b)$$

$$D_i^l \geq D_0^l y_i^l, \forall i \in I, \forall l \in \text{OAR}, \quad (2c)$$

$$\sum_{i \in I} y_i^l / N^l \leq P^l, \forall l \in \text{OAR}, \quad (2d)$$

$$D_i^T \leq d_i^T + M y_i^{TR}, \forall i \in T, \quad (2e)$$

$$D_i^T \geq d_i^T y_i^{TR}, \forall i \in T, \quad (2f)$$

$$\sum_{i \in T} y_i^{TR} / N^T \leq P^{TR}, \quad (2g)$$

$$D_i^T \leq d_i^T + M(1 - y_i^{TL}), \forall i \in T, \quad (2h)$$

$$D_i^T \geq d_i^T (1 - y_i^{TL}), \forall i \in T, \quad (2i)$$

$$\sum_{i \in T} y_i^{TL} / N^T \leq P^{TL}, \quad (2j)$$

$$D_i^T = \sum_j w_{ij}^T x_j, \forall i \in T \quad (2k)$$

$$D_i^l = \sum_j w_{ij}^l x_j, \forall i \in I, \forall l \in \text{OAR} \quad (2l)$$

$$y_i^l \in \{0, 1\}, \forall i \in I, \forall l \in \{\text{OAR}, T\}, \quad (2m)$$

$$y_i^{TL}, y_i^{TR} \in \{0, 1\}, \forall i \in T \quad (2n)$$

$$x_j \geq 0, \forall j \in J \quad (2o)$$

The d_i^T and d_i^T are the new decision variables for the dose-volume constraints of the cold and hot spots for tumors, respectively. We minimize their difference to ensure that we achieve the best balance. The objective (2a) is to minimize the difference between the dose of the cold and hot spots for tumors. Constraints (2b) and (2c) together are to decide whether the dose at a voxel i of a normal organ l exceeds the prescribed dose D_0 or not. If the dose at a voxel i of a normal organ l does not exceed the prescribed dose D_0 , then the corresponding binary variable y_i is 0. Otherwise, the corresponding binary variable y_i is 1. This is facilitated by a sufficiently large positive number M . Constraints (1d) are the dose-volume constraints for OARs with the help of the binary variable y_i , determined by (2b) and (2c). We limit the percentage of the organ l volumes with doses exceeding D_0^l to be less than or equal to P^l . Similarly, constraints (2e) to (2g) are the dose-volume constraints for tumor hot spots. We limit the percentage of the tumor volumes with doses exceeding d_i^T to be less than or equal to P^{TR} . Constraints (2h) to (2j) are for the dose-volume constraints for tumor cold spots. Equations (2k) and (2l) represent the total dose to each voxel from

all beamlets, and x_j are continuous variables representing the j th beamlet intensity (fluence), while the influence matrix, w_{ij} , is the dose contribution of the beamlet j of unit intensity at voxel i . Constraints (2m) and (2n) are the constraints for the binary decision variables. Constraints (1o) enforce the non-negativity of x variables.

3 | RESULTS

We have used Model 2 to determine the best parameters for all three cases. The optimal values derived are supplied to Model 1 to generate the plans for these three cases. Figure 1 shows the DVH results for the pediatric patient with prostate rhabdomyosarcoma, and Figs. 2 and 3 show the DVH results for the patients with H&N cancer. The corresponding DVH indices are also shown in Table 1 with the corresponding dose volume constraints.

From Figs. 1–3, by comparing DVH results of normal organs and tumors, we can see that our results provide comparable tumor dose coverage, tumor dose homogeneity and the maximum doses to OARs. At the same time, the mean doses of OARs are notably reduced. Another interesting observation is that the DVHs from the MIP method tend to be discontinuous, compared to the DVHs from the L-BFGS method.

Figure 4 shows the typical result of the dose distribution at one CT slice. Green line is for the CTV, cyan line is for the brainstem, red line is for the prescription iso-dose line (68 Gy) and dark blue line is for the iso-dose line of 54 Gy. We can clearly see the better brainstem protection from our model. However, small part of the tumor volume is underdosed. In the prostate case, after we enforced the dose-volume constraints in the MIP model, $D_{25\%}$ of the rectum and the bladder improved with dose-volume constraints. However, protection of the rectum and bladder came at the cost of tumor coverage. The $D_{95\%}$ of the tumors decreased in our result from MIP model than that from the quadratic programming method. For both H&N cases, interestingly both $D_{98\%}$ and $D_{2\%}$ – $D_{98\%}$ of the tumors are improved with better protections of all the normal tissues from our model.

4 | DISCUSSION

The MIP model with dose-volume constraints can generate treatment plans that meet all dose-volume requirements for OARs. The new model can generate IMPT treatment plans with comparable target dose coverage, target dose homogeneity, and the maximum dose to OARs compared to treatment plans from the conventional quadratic programming method without any tedious trial-and-error process. Some notable reduction in the mean doses of OARs is observed.

In addition, Model 2 is very effective to find the optimal parameters for obtaining the best balance between the tumor coverage and OAR sparing. The dose-volume constraints for the OARs are rigid especially for some important organs like spinal cord and brain stem.

In some scenarios tumor coverage may have to be sacrificed slightly to meet the dose volume constraints for these important organs.

To demonstrate our method, we have set constraint parameters to certain values to show that the plans generated from our proposed method meet our institution's dose volume constraints. We can certainly modify these parameters of dose volume constraints to achieve different results to protect normal organs better or relax the requirements for normal organs' protection to improve the tumor coverage. This is a medical decision and thus the user should be able to modify the corresponding parameters of dose volume constraints to achieve the desired plan depending on the patient specific clinical priority. Alternatively we can achieve the soft constraints in linear programming using so-called "fuzzy logic based optimization", which will generate the best normal tissue protection given the fixed tumor coverage without any tedious trial-and-error process. This is the ongoing research in our group. On the other hand, we can include soft constraints in the objective function in quadratic programming as proposed by Wu and Mohan³ This is the most popular way to implement the dose volume constraints in radiotherapy planning.

Generally, the maximum dose of normal tissue from the model with the dose-volume constraints is comparable to that from the quadratic programming method. However, the mean dose of some normal tissues is remarkably reduced (see Figs. 1–3). This might be important for the sparing of some important parallel organs such as parotids, oral cavity, total lung, etc. Interestingly from Figs. 1–3 the DVHs from the MIP method tend to be discontinuous, while the DVHs from the L-BFGS method are smooth. We speculate that this might be due to the hard constraints used in the MIP method and we believe that these insignificant discontinuities would not be clinically important.

Our MIP model only needs to run once to achieve the results, while quadratic programming method may need to run several times to get a better result (i.e., trial-and-error). We have done "trial-and-error" in quadratic programming to get the best results and compared the best results from quadratic programming with results from the MIP model. This shows the advantage of our model that the tedious "trial-and-error" process is not needed in our proposed method.

This advantage can be explained in more detail. The dose-volume constraints are hard in contrast to the soft constraints used in the conventional, heuristic methods. Our method uses an accurate model that produces a better result more efficiently. Our model ensures that the dose-volume constraints are rigidly met for both OARs and tumors. After solving our model, the optimal solution can satisfy all the dose-volume constraints because of these hard constraints. On the other hand, the result from any heuristic algorithm cannot guarantee that all the dose-volume constraints are met. Meeting the constraints can improve protection for OARs and thus can lead to better quality of life for patients. The proposed model can also inform the treatment planner that no optimal solution exists because of too stringent tumor dose constraints before we start to solve the problem. Heuristic methods do not have this capability. In addition, the iterations of the trial-and-error process used in heuristic

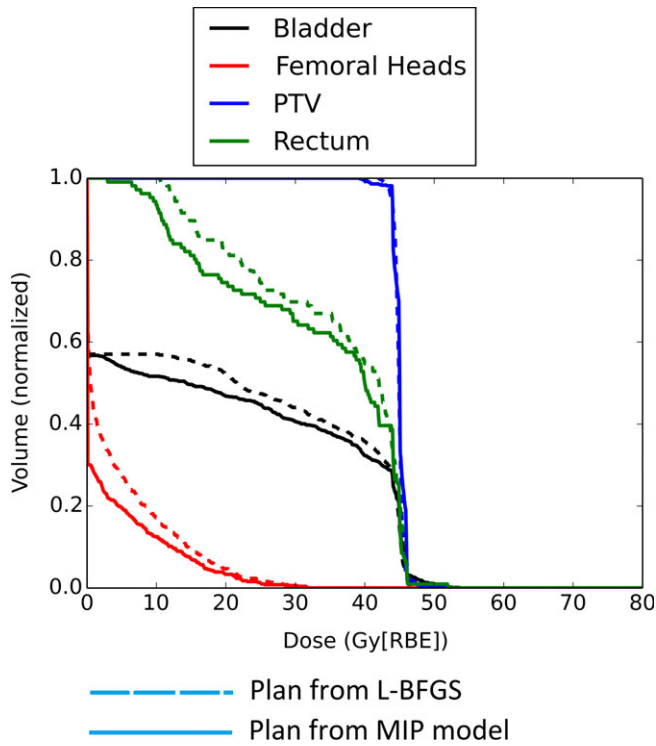


FIG. 1. DVH for the patient 1 with pediatric prostate rhabdomyosarcoma.

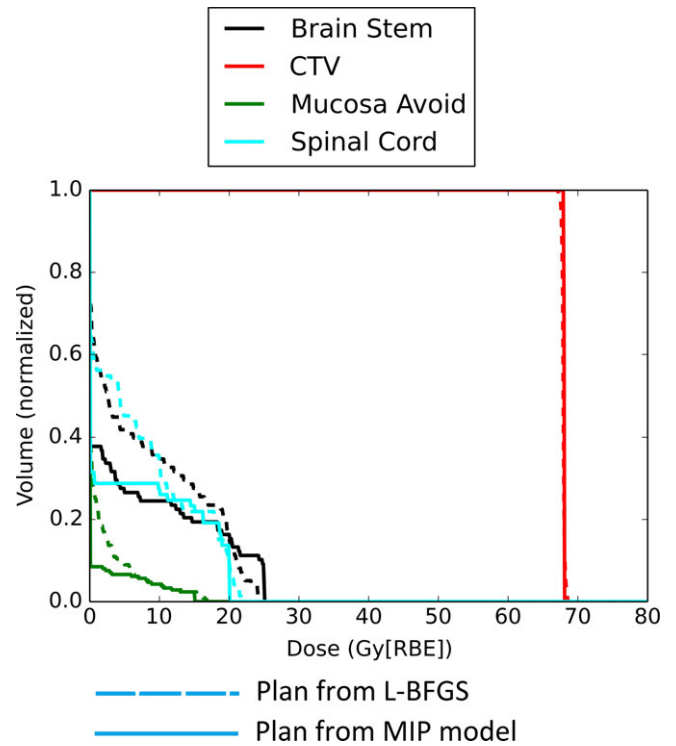


FIG. 3. DVH for patient 3 with H&N cancer.

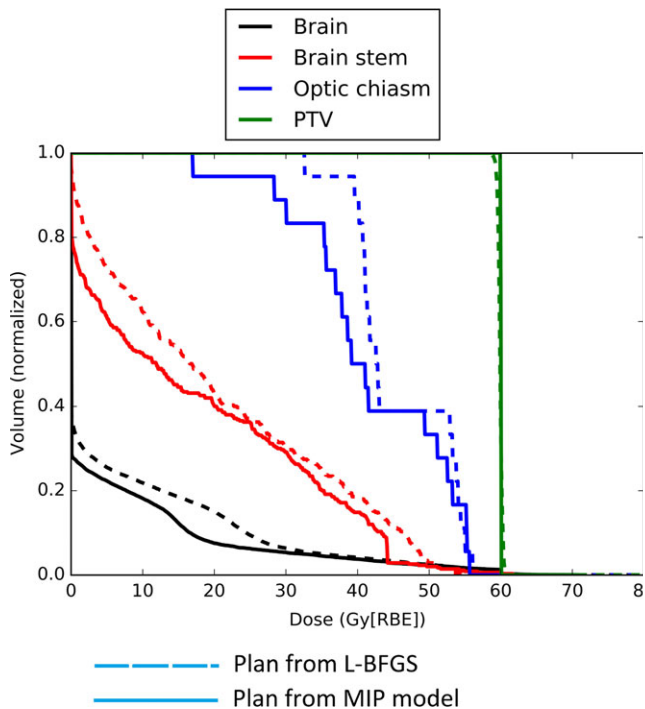


FIG. 2. DVH for patient 2 with H&N cancer.

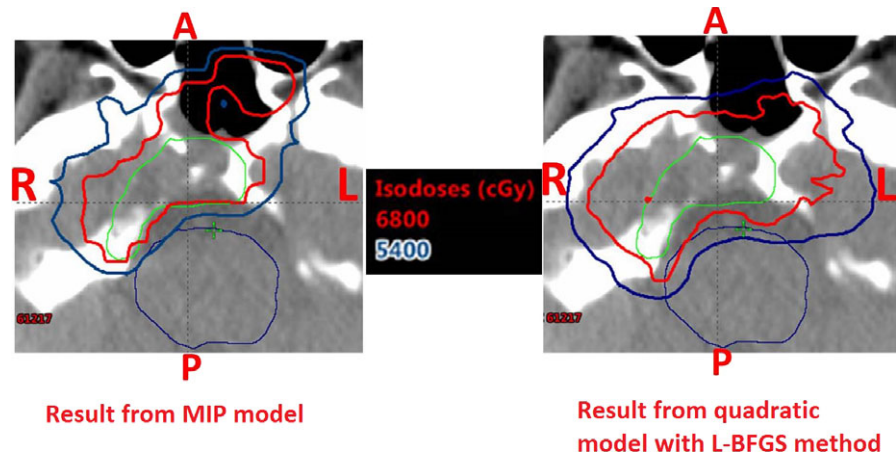
methods are time-consuming, and an optimal result is not guaranteed, while our second model can derive the best parameters for generating plans with the best balance between tumor coverage and protection of OARs.

We also compared the beam intensity results from our model with those from a quadratic programming model solved by the L-BFGS method. Interestingly, the number of beamlets with nonzero intensity was much lower in our linear model than in a nonlinear model, which is consistent with other reports.⁹ Fewer beamlets with nonzero intensity will improve the delivery efficiency, which is another advantage of our linear model.

Given the dose volume constraints of normal tissues, in some situations too stringent tumor dose constraints would make the Model 1 infeasible; while in some situations too loose tumor dose constraints would lead to undesired tumor dose distribution. Therefore it is important to set proper tumor dose constraint parameters in our MIP model. We used Model 2 to determine the optimal tumor dose constraint parameters that would achieve the best balance between tumor dose coverage and OAR sparing more efficiently. In all of our example cases, Model 2 allowed a much better balance than the conventional trial-and-error procedure. It also turns out that Model 2 always achieves satisfactory tumor dose distributions, which meet the tumor dose-volume constraints at least for the three clinical cases included in this study. We believe that the use of Model 2 would improve the quality of treatment plans.

Our model has certain limitations. First, a complex problem requires substantial time to reach a final, optimal solution, so we currently terminate the optimization if we have a feasible solution, whose objective value is within a certain gap tolerance compared to the optimal solution. In our examples, we can get a feasible solution with 50% gap within 2 hr in all cases (for the case of patient 3, we get the optimal result within 5 s). We believe that we can

FIG. 4. Dose distribution results at one CT slice for patient 2: MIP model with dose-volume constraints (left) and quadratic model solved by L-BFGS method (right). Green: Clinical target volume (CTV), Cyan: Brain stem. Red: prescription iso-dose line (68 Gy[RBE]), Dark Blue: iso-dose line of 54 Gy[RBE]. MIP model with dose-volume constraints generated a plan with much better brain stem protection at the cost of slight tumor underdosage.



significantly improve the efficiency of our current model by using either Benders' decomposition¹⁶ or parallel computing.

5 | CONCLUSION

We have proposed a model that can generate treatment plans meeting all the requirements for dose-volume constraints of OARs without any tedious trial-and-error process. This model has the potential to automatically generate IMPT plans with consistent plan quality among different treatment planners and across institutions and better protection for important parallel OARs in an effective way.

We are working toward developing more efficient algorithms that will allow us to use the model for patient population study. To incorporate robust optimization in linear programming for IMPT treatment planning is also another future research direction.

ACKNOWLEDGMENTS

This research was supported by the National Cancer Institute (NCI) Career Developmental Award K25CA168984, the Fraternal Order of Eagles Cancer Research Fund Career Development Award, the Lawrence W. and Marilyn W. Matteson Fund for Cancer Research, Mayo Arizona State University Seed Grant, and the Kemper Marley Foundation.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Suit HD. Protons to replace photons in external beam radiation therapy. *Clin Oncol (R Coll Radiol)*. 2003;15:S29–S31.
- Suit H, Urie M. Proton beams in radiation therapy. *J Natl Cancer Inst*. 1992;84:155–164.
- Wu Q, Mohan R. Algorithms and functionality of an intensity modulated radiotherapy optimization system. *Med Phys*. 2000;27:701–711.
- Falkinger M, Schell S, Müller J, Wilkens JJ. Prioritized optimization in intensity modulated proton therapy. *Z Med Phys*. 2012;22:21–28.
- Wu X, Zhu Y, Dai J, Wang Z. Selection and determination of beam weights based on genetic algorithms for conformal radiotherapy treatment planning. *Phys Med Biol*. 2000;45:2547.
- Xing L, Li JG, Donaldson S, Le QT, Boyer AL. Optimization of importance factors in inverse planning. *Phys Med Biol*. 1999;44:2525–21.
- Chen W, Craft D, Madden TM, Zhang K, Kooy HM, Herman GT. A fast optimization algorithm for multicriteria intensity modulated proton therapy planning. *Med Phys*. 2010;37:4938–4945.
- Cao W, Lim GJ, Lee A, et al. Uncertainty incorporated beam angle optimization for IMPT treatment planning. *Med Phys*. 2012;39:5248–5256.
- Zaghian M, Lim G, Liu W, Mohan R. An automatic approach for satisfying dose-volume constraints in linear fluence map optimization for IMPT. *J Cancer Ther*. 2014;5:198.
- Tuncel AT, Preciado F, Rardin RL, Langer M, Richard JPP. Strong valid inequalities for fluence map optimization problem under dose-volume restrictions. *Ann Oper Res*. 2012;196:819–840.
- Ferris MC, Meyer RR, D'Souza W. Radiation treatment planning: mixed integer programming formulations and approaches. *Handbook on Modelling for Discrete Optimization*. US: Springer, 2006:317–340.
- Rocha H, Dias JM, Ferreira BC, Lopes MC. Discretization of optimal beamlet intensities in IMRT: a binary integer programming approach. *Math Comput Model*. 2012;55:1969–1980.
- Romeijn HE, Ahuja RK, Dempsey JF, Kumar A. A new linear programming approach to radiation therapy treatment planning problems. *Oper Res*. 2006;54:201–216.
- Li Y, Zhu XR, Sahoo N, Anand A, Zhang X. Beyond Gaussians: a study of single-spot modeling for scanning proton dose calculation. *Phys Med Biol*. 2012;57:983.
- Liu W, Zhang X, Li Y, Mohan R. Robust optimization of intensity modulated proton therapy. *Med Phys*. 2012;39:1079–1091.
- Noonan F, Giglio RJ. Planning electric power generation: a nonlinear mixed integer model employing benders decomposition. *Manage Sci*. 1977;23:946–956.