



Costlets: A Generalized Approach to Cost Functions for Automated Optimization of IMRT Treatment Plans

MARC L. KESSLER*

DANIEL L. MCSHAN

Department of Radiation Oncology, The University of Michigan Medical School, Ann Arbor, MI 48109-0010

email: mkessler@umich.edu

MARINA A. EPELMAN

Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, MI 48109-2117

KAREN A. VINEBERG

AVRAHAM EISBRUCH

THEODORE S. LAWRENCE

BENEDICK. A. FRAASS

Department of Radiation Oncology, The University of Michigan Medical School, Ann Arbor, MI 48109-0010

Received 12 March, 2003; Revised 6 May, 2004

Abstract. We present the creation and use of a generalized cost function methodology based on costlets for automated optimization for conformal and intensity modulated radiotherapy treatment plans. In our approach, cost functions are created by combining clinically relevant “costlets”. Each costlet is created by the user, using an “evaluator” of the plan or dose distribution which is incorporated into a function or “modifier” to create an individual costlet. Dose statistics, dose-volume points, biological model results, non-dosimetric parameters, and any other information can be converted into a costlet. A wide variety of different types of costlets can be used concurrently. Individual costlet changes affect not only the results for that structure, but also all the other structures in the plan (e.g., a change in a normal tissue costlet can have large effects on target volume results as well as the normal tissue). Effective cost functions can be created from combinations of dose-based costlets, dose-volume costlets, biological model costlets, and other parameters. Generalized cost functions based on costlets have been demonstrated, and show potential for allowing input of numerous clinical issues into the optimization process, thereby helping to achieve clinically useful optimized plans. In this paper, we describe and illustrate the use of the costlets in an automated planning system developed and used clinically at the University of Michigan Medical Center. We place particular emphasis on the flexibility of the system, and its ability to discover a variety of plans making various trade-offs between clinical goals of the treatment that may be difficult to meet simultaneously.

Keywords: optimization, cancer, radiation therapy, mathematical programming, intensity modulated radiation therapy, treatment planning

1. Introduction

Treatment planning for external-beam radiation therapy is inherently an optimization problem. Because it is not possible to irradiate a cancerous tissue without also partially irradiating

*Corresponding author.

surrounding healthy tissues, tradeoffs are made to balance the competing issues of controlling or reversing the growth of the tumor and reducing the probability of radiation-induced side-effects to the healthy tissues. In order to reduce the doses to healthy tissues while still delivering a high uniform dose of radiation to a tumor, *conformal radiotherapy* makes use of a set of radiation beams produced by a gantry mounted linear accelerator (Greene, 1985) and delivered from different angles so that all fields overlap at the tumor. Each beam is independently shaped using a multileaf collimator device (Brahme, 1987) to help shield healthy tissue. In conventional radiotherapy, the radiation beams have either uniform or linearly varying intensity (created using metal filters). When designing a radiotherapy treatment plan for a patient, the *treatment planner* selects the number of beams used, the direction and shape of each beam and the relative intensities of the beams. A manually designed treatment plan typically consists of between 2 and 9 beams. The planner iterates through various candidate treatment plans, using both visual and numerical tools to help evaluate the relative merits of the different tradeoffs made in each plan. By increasing the number of the beams used, higher doses received by tissues outside the tumor can be "spread out" and reduced. This also, however, increases the complexity of the manual treatment planning process.

Another approach to reducing doses to healthy tissues while still delivering a therapeutic dose to the tumor is to vary the intensity across each beam in a nonlinear fashion. A number of investigators have reported on the theoretical advantages of this approach (Brahme, 1987, 1988, 1993; Webb, 1989, 1991; Bortfeld, 1993, 1994; Mohan, 1994). Several different collimation and modulation systems currently available are capable of delivering radiation beams with such intensity patterns. The general term for this delivery technique is *intensity modulated radiation therapy*, or IMRT. Conceptually, this approach can be represented by dividing each shaped beam into a larger number of smaller "beamlets," whose intensity, or *weight*, can be set independently.

Because of the large number of degrees of freedom available with IMRT (typically hundreds or thousands of beamlets are used to define the intensity variation for each treatment beam), manual planning of IMRT plans is not feasible. This has accelerated the development and deployment of treatment planning systems that automate the process of generating and evaluating alternative treatment plans. In these systems, instead of manipulating the beam directions, shapes and weights, the treatment planner manipulates a set of numerical criteria which represent the ultimate clinical goals of the treatment; these numerical criteria are used by the planning system to guide the automated search for different candidate plans and evaluate their relative merits and acceptability. The common term used for this approach to treatment design is *inverse treatment planning* to differentiate it from the manual, or forward, process of treatment planning.

Inverse treatment planning systems consist of several basic components, namely, patient-specific anatomic models, a dose calculation model, a search algorithm used to generate candidate treatment plans, and a mathematical model used to evaluate the relative value and acceptability of different candidate treatment plans. While there are still considerable research and development efforts underway to improve each of these components, the first two are fairly well established and common to both forward and inverse treatment planning systems. All modern treatment planning systems use volumetric medical imaging data and

image processing and computer graphics tools to construct and represent three-dimensional patient anatomy. X-ray computed tomography (CT) is the most common modality used to acquire the necessary anatomic information, though most systems now support the use of magnetic resonance (MR), nuclear medicine and ultrasound image data and provide tools to geometrically register and combine the information from these complementary modalities to construct patient models (Kessler, 1991; Ten Haken, 1992; Meyer, 1997).

In contrast to models for defining and representing patient anatomy and dose calculation models, prediction and modeling of the clinical outcome from a delivered dose distribution is still very much a research topic. A major focus of this area of research is the study of effects of non-uniform or partial-volume irradiation of healthy, or *normal*, tissues. Clearly, the ability to quantify these effects is necessary to design and implement a definitive mathematical model to evaluate and compare acceptability and relative merits of candidate plans within an automated planning system. However, in general, the impact of different dose distributions delivered to various healthy tissues is not sufficiently well understood to give rise to a universally acceptable mathematical model. Moreover, the flexibility of IMRT delivery, which can produce a nearly infinite variation of doses delivered to normal tissues while delivering fairly uniform high doses to tumors, complicates the discovery of the necessary dose-volume-complication relationships.

A variety of inverse planning systems currently exist, both commercially available and those developed at academic centers (Alber, 2001; Bortfeld, 1994; Brahme 1993; Carol, 1994; Langer, 1998; Holmes, 1995; Mohan, 1992; Spirou, 1998; Webb, 1991; Xing 1999). While the various systems may use different approaches to generate and compare candidate treatment plans, most are based on the use of a *single* type of mathematical model and a *single* search algorithm. Most of the early inverse planning systems used simple dose-based cost functions such as least-squares dose to one or more anatomic structures, and employed gradient-based search algorithms which perform well for quadratic cost functions (Bortfeld, 1999; Webb, 1989) although other more complicated functions were proposed (Agren, 1990; Bortfeld, 1988; Kallman, 1988; Mohan, 1996; Niemierko, 1992; Wang, 1995).

More recently, a number of increasingly sophisticated mathematical programming models have been proposed for the inverse treatment planning problem and some of these have made their way into clinical use. A survey by Shepard, et al. (1999) discusses several possible linear, nonlinear, and mixed integer optimization models of the treatment planning problem. For example, linear programming models were considered in Holder (2001, 2003) and Romeijn (2003), integer programming formulations were presented in Ferris (2002), Lee (2003), and Preciado (2004) while Lim (2002) described nonlinear models. In other work, the inverse treatment planning problem is treated as a feasibility problem (Xiao, 2003; Michalski 2003), as well as a linear or nonlinear multicriteria optimization problem (Hamacher, 2002; Kyfer, 2003). While these references are not a comprehensive overview of the literature on mathematical models developed for treatment planning, they help demonstrate the diversity of models being investigated.

Despite the differences in specific mathematical techniques employed, a typical optimization model for inverse treatment planning imposes a number of constraints on the decision variables (e.g., beamlet weights) by restricting the doses delivered to the tumor

and normal tissue, as prescribed by the physician (such as minimum and maximum doses on various anatomic structures), and gives an objective or *cost* function to be optimized subject to those constraints (e.g., the total dose received by healthy tissues, which should be minimized). However, because of the complex nature of how a radiation beam deposits dose in the patient, treatment planning is not a straightforward constrained optimization problem. Since the ideal treatment—one that delivers a high uniform dose of radiation to the tumor and essentially no radiation to other tissues—is impossible to deliver, the clinician usually has to accept some compromises, achieving an appropriate tradeoff between the contradicting goals above. A very important issue in the use of any planning system is how to make tradeoffs between dose to the cancerous tissues, or *target volumes*, (to increase tumor control probability) and dose to normal tissues (to decrease normal tissue complication probabilities). In particular, it is critical to be able to evaluate and compare plans that make different tradeoffs between competing goals. As noted above, most inverse planning systems use a single prespecified type of objective, or cost, function. In contrast, in this paper we describe an extremely flexible planning system which allows the planner to construct and optimize a diverse family of cost functions and systematically compare the resulting treatment plans.

From a mathematical programming standpoint, our approach is essentially a penalty method for constrained optimization (see for example, Bazarra, 1993). That is, we construct the cost function to penalize plans that will result in clinically undesirable outcomes, and use a search algorithm to find a solution (i.e., the optimal set of beamlet weights) that minimizes the overall penalty. Our cost function is based on the use of *costlets*: individual “pieces” of the cost function which describe, for a particular structure, the desired treatment goal as well as the penalties for not achieving that goal. By combining costlets, it is possible to develop complex yet understandable and clinically relevant overall cost functions for use in clinical treatment planning protocols and optimization studies. Our methodology can help the treatment planner to incorporate much of the true complexity of the radiotherapy plan design into an automated planning system.

In this paper, we report the motivation, design and use of costlet-based plan scoring capabilities. We illustrate the costlet-based methodology using a clinical example involving treatment plan design for a head-and-neck cancer. Clinical optimization studies using this costlet-based approach have been reported for various clinical sites, including brain (Vineberg, 1999), head-and-neck (Vineberg, 2000, 2002), prostate (Damen, 2001), lung (Seppenwoolde, 2002), and breast (Krueger, 2003), and other studies are ongoing. The contributions of this paper are two-fold. On the one hand, we use traditional optimization methodology to solve problems arising from a novel and complex application. On the other hand, we address the medically relevant issue of formulating optimization problems to aid in design of plans that aim to satisfy different, often conflicting, clinical goals.

The remainder of this paper is organized as follows. In Section 2 we outline the preparatory steps and calculations that need to be carried out prior to the construction of the optimization problem for treatment plan design. In Section 3, we provide a detailed description of the costlet-based model. In Section 4, we give an example of the planning process for a specific head-and-neck case, and compare the plans developed when using various cost functions. Section 5 contains a discussion and concluding remarks.

2. Treatment planning preparation

The treatment planning process begins by creating 3-D models of the patient anatomy in the treatment position. X-ray CT is the primary imaging modality used for this, although magnetic resonance and nuclear medicine data might also be used to help improve localization of the disease and critical structures. Using this image data, the clinician outlines the regions to treat and the relevant healthy tissues to avoid (Figure 1). While some amount of automation is possible, manual outlining is still a labor intensive process. In most cases, the clinician manually defines a *gross tumor volume* (GTV), which is the visible gross disease, and also a *clinical tumor volume* (CTV), which is either an isotropic or anisotropic expansion of the GTV that includes areas of probable microscopic spread that are not detected by imaging. In some cases, the CTV is defined directly based on known patterns of disease spread. The CTV is then expanded to account for estimated physiological movement such as breathing and small variations in patient position over the course of treatment (a typical course of radiotherapy is delivered in daily fractions over a four to six week period). This expansion is called the *planning target volume* (PTV) and it is designed so that radiation beams shaped to this volume will encompass the actual tumor volume 100% of the time (ICRU, 1993). While most treatments involve a single PTV, it is possible to have several different target volumes that may be prescribed to different dose levels reflecting different levels of risk.

Using the 3-D anatomic model as a guide, the treatment planner selects the number and locations of the radiation beams to be used in the treatment. In conformal therapy using uniform or *flat* fields, the arrangement of beams is carefully selected to minimize the overlap with healthy tissues. Most IMRT treatments involve simple beam arrangements of 5, 7 or 9 coplanar beams equally spaced around the patient, since the intensity pattern of each beam varies by design. Odd numbers of beams are used so that no two beams directly oppose each other. Each beam is subdivided into a rectilinear grid of beamlets. Typical size beamlets are 1 cm \times 1 cm but may be as small as 0.2 cm \times 0.5 cm. Beamlets which do not intersect a PTV (plus some margin) are not considered in the optimization (weight is fixed at zero) and are not displayed. Figure 2(a) shows a beam arrangement for a case involving a tumor of the head-and-neck. The various level of shading of the beamlets in the figure represents various levels of intensity, or weights, of the beamlets used for treatment (brighter = higher intensity). In some of the beams, a darker region can be seen near the center of the intensity map. The beamlets in this part of the intensity map intersect the spinal cord and have been kept to low intensities by the optimization process. Lower intensity beamlets that intersect the parotid glands can also be seen.

In our current implementation the number and orientation of the beams are held fixed. While there can be advantages to varying these parameters during the optimization process, each new beam or beam direction requires a computationally expensive dose calculation (Bortfeld, 1993; Rowbottom, 2001; Meedt, 2003). With current hardware configurations available in clinical settings, this would add significant time (hours to days) to the overall optimization process and reduce the clinical utility of the system, especially when multiple optimizations representing different clinical tradeoffs may be required. By using a sufficient number of beams, the importance of beam orientation is usually reduced.

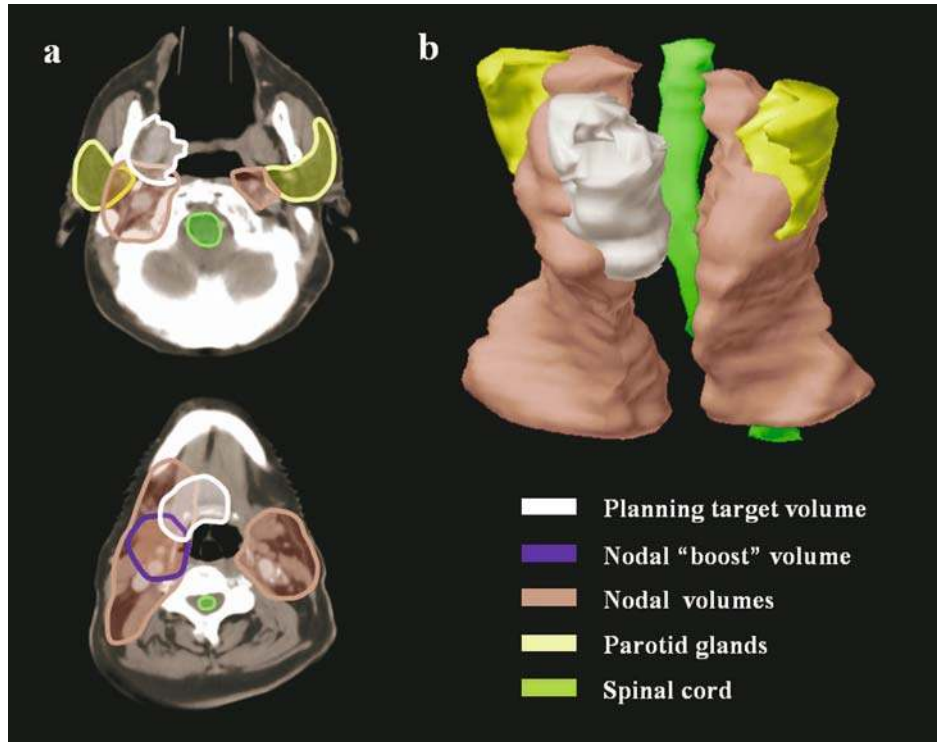


Figure 1. (a) CT images at different levels (oral cavity and bottom of the jaw) with outlines of the different structures to be considered during optimization. (b) 3-D models derived from the outlines.

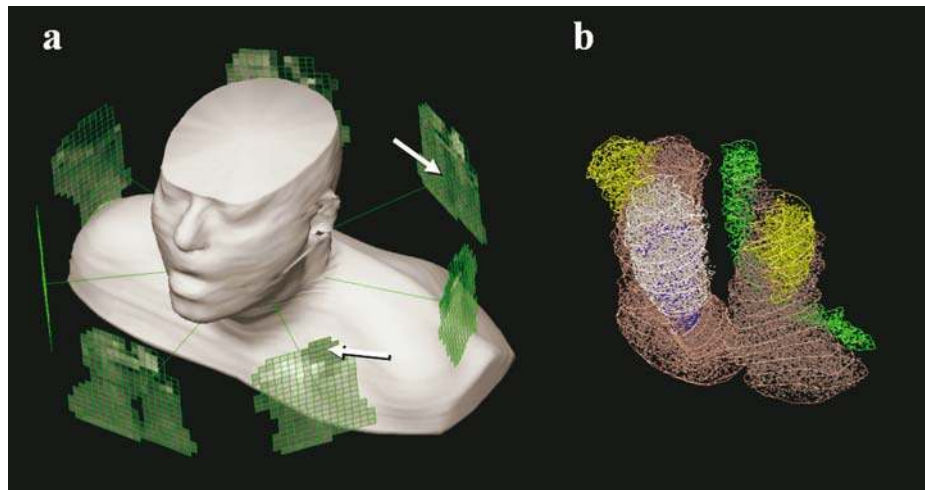


Figure 2. (a) 7-field beam arrangement for head-and-neck cancer patient, (b) Dose calculation points sampled from structures.

Once the patient anatomy and beams are defined, the elemental dose (unit weight) to a set of sample points for each structure from each beamlet is computed. The locations of the sample points are determined using a pseudo-random algorithm that attempts to place points uniformly throughout the 3D structure (Press, 1992). The density of sample points is chosen to adequately represent both the shape of the particular structure and the expected dose variation across the structure. The volume of each sample point is determined by the volume of the structure divided by the number of sample points per structure. Typical sample volume sizes range from 0.001 cc to 1.0 cc. For example, a PTV with a volume of 100 cc might consist of 10,000 points. For the head-and-neck example case a total of 33,117 points were used to represent 7 structures. It is also common to create sample points for composite or partial structures using Boolean operations. An example is the *normal brain* which might be constructed by removing the PTV from the whole brain. Figure 2(b) illustrates the point-based representation of the patient anatomy for the example case described in Section 4.

The dose calculation engine used to compute the elemental doses in this work is part of our clinical 3-D treatment planning system, U-MPlan, which was developed in-house and has been in clinical use for over 18 years (Fraass, 1987a, 1987b; McShan, 1987, 1990). This system runs on OpenVMS-based workstations and servers (Hewlett-Packard, Palo Alto, CA) using software written in FORTRAN and C. Three-dimensional photon dose calculations for the beamlet IMRT fields are performed using a convolution/superposition algorithm based on an algorithm originally provided by Mackie and colleagues (Mackie, 1985). Since only the beamlet weights are manipulated during each iteration of the optimization process, it is possible to express the dose delivered by any set of intensity maps as a sum of the contributions of the individual beamlets, making it unnecessary to calculate the elemental doses at each iteration.

3. Costlets and construction of the objective function

Once the initial dose calculation to the sampled points from each defined beam has been performed, an objective or cost function, i.e., a mathematical formulation of the plan design problem, is constructed. Overall cost functions are built up using three components: *evaluators*, *modifiers* and *costlets*. In this section we provide a general description of each component and its contribution to the overall cost function. A simple example is also provided. In the following section we provide examples of cost functions designed to address the planning problem for a particular patient and compare the resulting solutions.

3.1 Evaluators, modifiers and costlets

The first step in the construction of an objective function consists of interactively choosing one or more *evaluators* from a predefined list. Broadly defined, an evaluator is a function $E(\cdot)$ mapping the calculated dose distribution (or some other information, e.g., the number of beamlets with nonzero intensity) into a scalar; thus an evaluator is a simple numeric

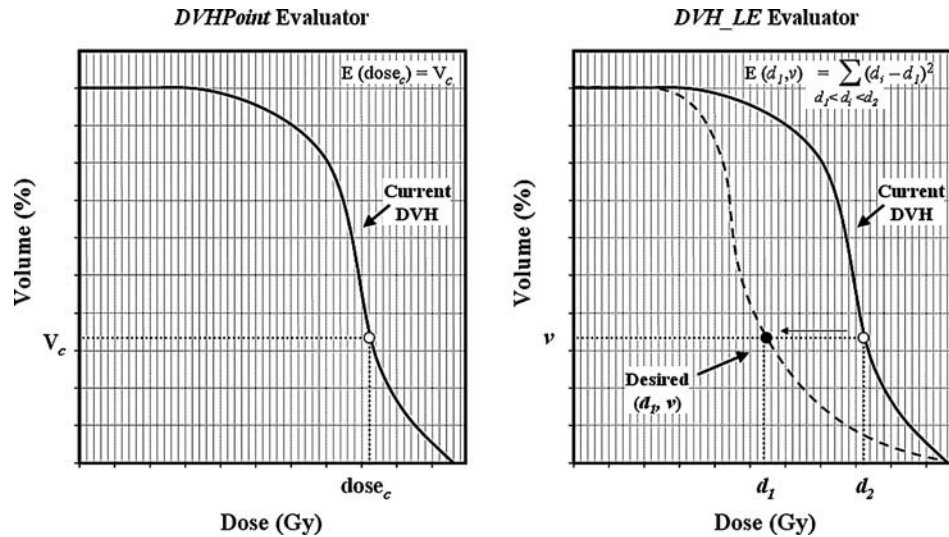


Figure 3. Example evaluators.

metric of a plan. Each evaluator is designed so that its value quantifies and conveys a relevant feature of the plan.

Simple evaluators include dose statistics such as the minimum, maximum, mean (average), or variance of the dose to a structure. The meanings of these evaluators are easy to understand and allow the planner and physician to relate the optimization system results to their clinical experience; they help to provide an intuitive understanding of the trade-offs made during optimization. More complex evaluators can also be included. Evaluators that reflect some aspect of the dose-volume relationship for a particular structure are common. Specification of one or more parameters is often required to define such evaluators. Figure 3 illustrates a *DVHpoint* evaluator with parameter dose_c , defined as the fraction of the volume of the structure receiving a dose larger than dose_c , and the *DVH_LE* with parameters (d, v) which is a desired dose-volume endpoint for a particular structure (see next section for a detailed explanation).

Simple biological model functions such as normal tissue complication probability (NTCP) (Lyman, 1985), tumor control probability (TCP) (Goitein, 1987) and equivalent uniform dose (EUD) (Niemierko, 1997) may also be used as evaluators since they specify, for each dose level d , what fraction of the volume of the structure receives a dose greater than or equal to d . For a description of a multi-criteria modeling paradigm using biological model functions see Küfer (2003). Table 1 lists some common evaluators that have proven useful in both inverse and forward 3-D planning experience.

For each evaluator chosen, a *modifier* is used to convert the evaluator into a *costlet*. A modifier is a function of one variable, usually restricted to have nonnegative values. Each resulting costlet c has the form $c = f(E)$, where E is the evaluator used, and f is the modifier chosen for this evaluator. Each costlet is meant to convey a specific clinical goal

Table 1. Typical evaluators.

Name	Parameters	Description/definition	Example of output
Min		Minimum dose to structure	Minimum dose to points in PTV
Max		Minimum dose to structure	Minimum dose to points in the spinal cord
Mean		Mean (average) dose to structure	Average dose to PTV
DVHpoint	D_t : Target dose to the structure	% vol. of the structure above specified dose D_t	% vol. of cord receiving dose greater than 50 Gy
nDVHpoint	D_t : Target dose to the structure	% vol. of the structure below specified dose D_t	% vol. of PTV receiving dose less than 80 Gy
Threshold	D_t : Upper dose threshold p : Power	Sum, over all points in the structure, of the positive part of the difference between dose to the point and D_t raised to the power p	$\sum_{i \in p \text{ in cord}} (\max(d_i - 50, 0))^4$ here $D_t = 50, p = 4$
nThreshold	D_t : Lower dose threshold p : Power	Similar to Threshold, with penalty incurred for delivering dose below D_t	$\sum_{i \in p \text{ in PTV}} (\max(63 - d_i, 0))^4$ here $D_t = 63, p = 4$
Tolerance	D_t : Target dose to the structure l : % violation, p : power	Similar to Threshold, with penalty incurred for delivering dose outside the range $D_t \pm l \bullet 100\%$	$\sum_{i \in p \text{ in PTV}} (\max(d_i - 105, 95 - d_i, 0))^2$ here $D_t = 100, l = 5, p = 2$
DoseLSQ	D_t : Target dose to the structure	Sum, over all points in the structure, of squares of difference between dose to the point and D_t	$\sum_{i \in p \text{ in PTV}} (d_i - 100)^2$ here $D_t = 100$
NTCP	m, n, TD_{50} (see references)	Normal tissue complication probability (Lyman, 1985, 1992).	NTCP for lung
EUD	a : volume parameter	Equivalent uniform dose (Niemierko, 1997). Uniform dose to the organ that leads to the same effect as non-uniform dose.	$(\frac{1}{n \text{ pts}} \sum_{i \in p \text{ in lung}} d_i^4)^{1/4}$ here $a = 4$

(continued on next page.)

Table 1. (continued).

Name	Parameters	Description/definition	Example of output
V_{eff}	reference dose, n (see reference)	Effective volume (Kutcher, 1989). Volume of organ that if treated with a uniform dose leads to the same effect as non-uniform dose.	V_{eff} for right lung
# segments		Number of treatment segments, estimated from plan description	
T_x time est		Time required to deliver treatment, estimated from plan description	

of the physician with respect to an individual clinical issue via a quantitative “cost,” or “penalty,” incurred by a plan and resulting dose distribution. Therefore, the form of the modifier chosen for an evaluator is meant to convey a particular nature of dependence of the penalty on the property of the plan expressed in the evaluator. In particular, a combination of an evaluator and modifier for each costlet should be chosen so that higher values of the resulting costlet correspond to less desirable outcomes with respect to the clinical goal that the costlet is meant to convey. For example, if the clinical goal at hand is to make the mean dose to a certain structure (e.g., the liver) as small as possible, a possible costlet is $c_1 = \max(0, E_1)$, where E_1 is the “Mean” evaluator for the structure. If, on the other hand, the clinical goal is to keep the mean dose to PTV within the prescribed range (say, between 85 and 95 Gy), then the costlet chosen to reflect this goal might be $c_2 = \max(E_2 - 95, 0) + (\max(85 - E_2, 0))^2$, where E_2 is the “Mean” evaluator for PTV. Notice that this costlet imposes a penalty for both underdosing and overdosing the structure, with a steeper penalty for underdosing than for overdosing. As the example demonstrates, the relative importance of different plan properties expressed through evaluators can be reflected in the choice of modifiers.

Typically, modifiers with simple functional forms are used when there is no evidence to defend any particular costlet shape. Linear, quadratic and higher power functions of evaluators, as well as more complex threshold-based functions, can be used. Figure 4 illustrates two types of modifiers which can be used for nonnegative evaluators. In the first example, the modifier is a piecewise linear function, which uses a small slope for small values of the evaluator and a larger slope for the values of evaluator above E_b (modifiers of this type are referred to as “3ptCeiling”). The second example modifier creates a costlet with no penalty until the threshold value E_t is reached, and then uses a high power (here, power 6) of the difference between the evaluator and the threshold value E_t . Other example modifiers are listed in Table 2, expressed as functions of the value of a generic evaluator, E .

Table 2. Typical modifiers.

Name	Parameters	Functional form
Passthrough		E
Power	A : Weight, p : Power	$A \bullet E^p$
FloorPower	A : Weight, F : Floor, p : Power	$A(\max(0, E - F))^p$
CeilingPower	A : Weight, C : Ceiling, p : Power	$A(\max(0, C - E))^p$
EqPower	A : Weight, E_t : Target value, p : power	$A \bullet E - E_t ^p$
3ptCelling	E_t : Target value, E_b : Breakpoint ($E_b > E_t$), A_1 : Slope 1, A_2 : Slope 2 ($A_2 > A_1$)	$A_1 \bullet \max(0, E - E_t) + A_2 \bullet \max(0, E - E_b)$

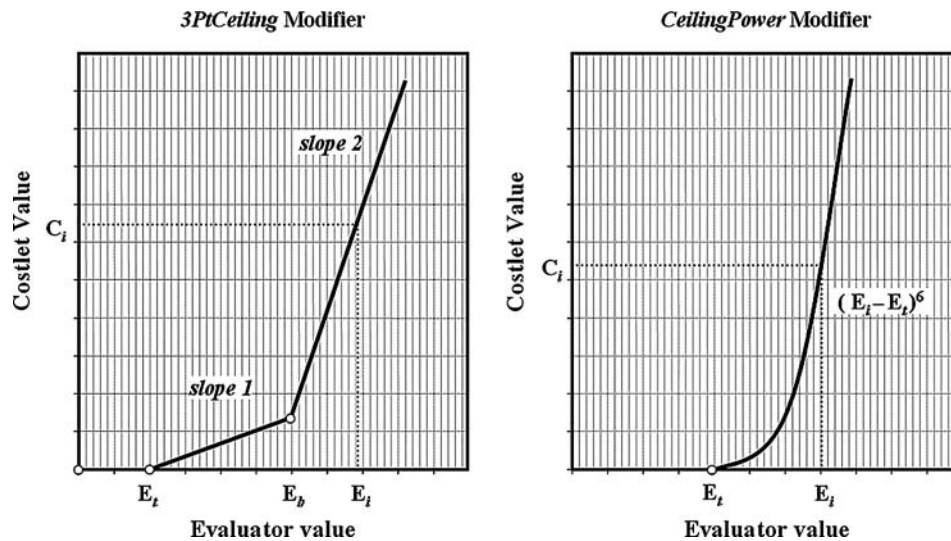


Figure 4. Example modifiers.

3.2 Overall cost function

The overall cost function $C(\cdot)$ is obtained by combining all the defined costlets. Specifically,

$$C = F(c_1, c_2, \dots, c_I),$$

where $c_i, i = 1, \dots, I$, are the costlets created as described earlier in this section, and the form of the function $F(\cdot)$ reflects the overall optimization strategy. The most obvious way to combine costlets into a penalty function, and the one used in the examples presented in this paper, is to use a weighted sum of costlets: $C = \sum_i w_i c_i$, where $w_i, i = 1, \dots, I$, are nonnegative weights used to further emphasize clinical tradeoffs between individual costlets. Other forms of $F(\cdot)$ implemented include the product of costlets, the probability of uncomplicated control (P^+) (Kallman, 1992) as well as a method which combines the sum and product methods.

ORGAN: LT_PAROTID - Cost: 0.00		name	min	max	mean	rms	t	c	color	cost
N	ON	CORD	0.00	45.17	15.06	10.45	0	0	blue	0.000
N	ON	RT_PAROTID	1.92	71.11	26.62	24.10	0	0	green	0.379
N	ON	LT_PAROTID	1.49	65.16	26.03	23.47	0	0	yellow	0.000
T	ON	PTV1	62.68	81.03	72.02	2.03	0	0	blue	0.116
T	ON	RTN	23.77	76.94	66.01	3.30	0	0	orange	5.246
T	ON	LTN	52.85	69.26	62.00	1.56	0	0	purple	0.036
T	ON	RT_BST	69.15	77.05	72.00	1.40	0	0	red	0.002

Figure 5. Structure browser with dose statistics and costs.

Structure: LT_PAROTID	num	t	status	eval	costlet
0	0	ON	DVH_LE	0.000	Passthrough
1	0	ON	Mean	26.000	Ceilingpow
2	0	OFF			
3	0	OFF			
4	0	OFF			
5	0	OFF			

Evaluators applied
(to left parotid)

<input type="radio"/> DVHpoint	<input type="radio"/> NDVHpoint	<input type="radio"/> Threshold	<input type="radio"/> NThreshold	<input type="radio"/> Tolerance
<input type="radio"/> StDev	<input type="radio"/> Lyman	<input type="radio"/> Veff	<input checked="" type="radio"/> Mean	<input type="radio"/> Min
<input type="radio"/> Max	<input type="radio"/> DoseLSQ	<input type="radio"/> MGHTCP	<input type="radio"/> UMTCP	<input type="radio"/> Conformity
<input type="radio"/> Homogeneity	<input type="radio"/> DVHRange	<input type="radio"/> MaxMean	<input type="radio"/> EUD	<input type="radio"/> FuncEUD
<input type="radio"/> Seriality	<input type="radio"/> DVH_EQ	<input type="radio"/> DVH_LE	<input type="radio"/> DVH_GE	<input type="radio"/> TolMean
<input type="radio"/> NegMean	<input type="radio"/> LymanP	<input type="radio"/> XSVolP		

Evaluator choice list

<input type="radio"/> Passthrough	<input type="radio"/> Floor	<input type="radio"/> EQ	<input type="radio"/> Ceiling	<input type="radio"/> 3ptCeiling
<input type="radio"/> EQpow	<input type="radio"/> Floorpow	<input checked="" type="radio"/> Ceilingpow	<input type="radio"/> OneOver	<input type="radio"/> Log

Modifier choice list

Ceiling: 26 Power: 2 Scale: 1

Figure 6. Costlet browser.

In our optimization system, costlets are chosen and reviewed using a *costlet browser*. Screenshots of the system interface are shown in Figures 5–7. First, one selects the desired anatomical structure from the structure browser and then the individual costlet. The evaluator and its parameters (if any) are selected. The modifier is then selected, and any modifier parameter(s) set, including the relative weight of the costlet. Various other controls for cost function template creation and selection are also included. Our implementation also allows one to label a cost as a true constraint which means that only plans which completely satisfy the costlet (value = 0) are considered viable.

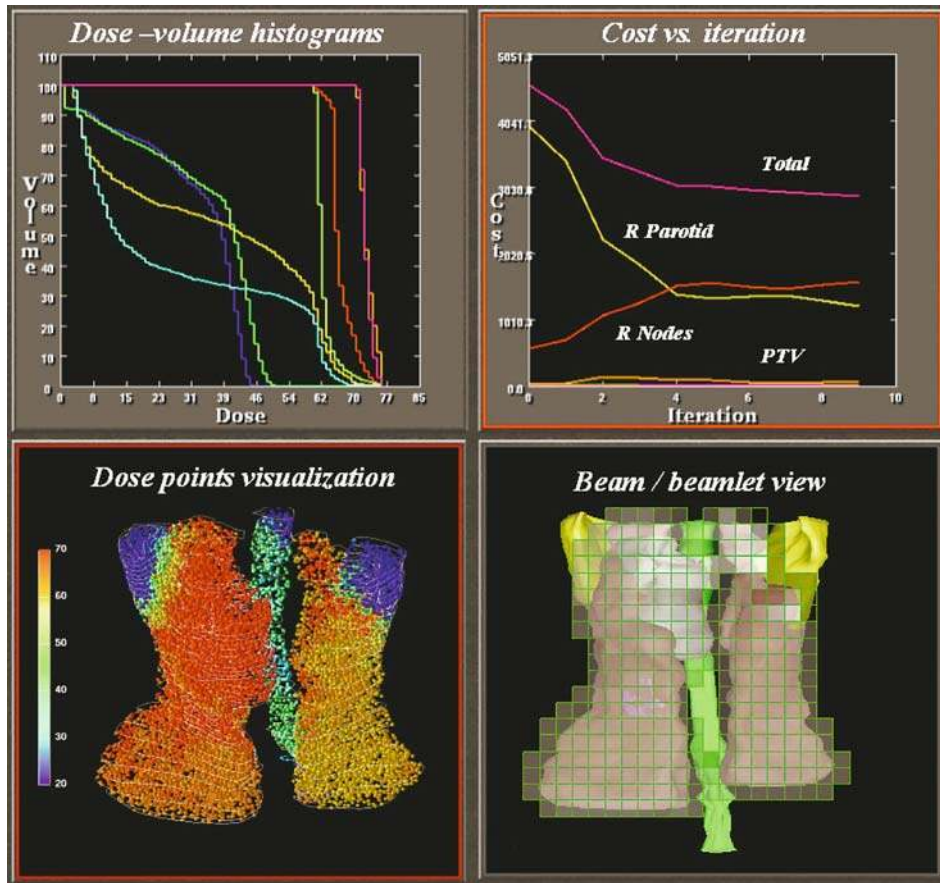


Figure 7. Data visualization interface. Graphics are updated during plan optimization.

3.3 Designing a cost function

With the costlet browser feature implemented, the mechanics of creating a cost function for any particular clinical situation become straightforward. The time-consuming part of the planning process is the creation of an *appropriate* cost function whose minimization will allow us to achieve the desired overall results.

The most important part of the process is the first step, a non-technical decision process which is performed mainly by the physician. The physician must identify clinical issues that should influence the optimization and prioritize their importance. The ranking of these clinical issues is critical, since some consequences may be extremely important to avoid (e.g., spinal cord injury), while others are clinically less important. An example of a clinical prioritization list for a plan for a head-and-neck case is shown below:

1. Do not give the spinal cord more than 45 Gy.
2. Do not underdose targets.
3. Do not overdose targets.
4. Give a mean dose of 60 Gy to targets.
5. Give as little dose as possible to the spinal cord.
6. Give as little dose as possible to other normal tissues.

Once all the clinical issues have been stated and prioritized, the planner must define the cost function which will attempt to reflect the specified clinical goals. For the current example priority list, at least 6 different groups of costlets would be developed, to incorporate each of the six clinical issues listed. The shape of the costlet (i.e., the functional form of the modifier) is often determined based on the priority assigned to the particular clinical property of the plan expressed by the evaluator used in the costlet (though the weight of the costlet in the objective function can also, in part, reflect its priority). For the example priority list above, the first costlet might be a dose-based costlet which penalizes points in the spinal cord whose dose is larger than 45 Gy using a high power of the dose above the threshold (power 10, for example). Such costlet will have the form

$$c = \sum_{j \in SC} \max(d_j - 45, 0)^{10},$$

where SC denotes the set of all points in the spinal cord identified for planning purposes. Here, for each point, we used evaluator *Threshold* with parameter 45, and modifier *Power* with parameters 1 and 10. This setting of the parameter in the modifier makes for a very steep penalty for exceeding the dose of 45 Gy to any point of the spinal cord, which will prevent high spinal cord doses. A similar costlet, but with a different power parameter in the modifier, would be appropriate for the lowest priority item on the above list (#6):

$$c = \sum_{j \in NT} \max(d_j, 0)^2,$$

where NT denotes the set of all points in the normal tissues. The remaining costlets are defined and prioritized using functional forms and parameters of evaluators and modifiers, so that the relative penalty induced by the individual costlets varies in concert with the priorities of the corresponding clinical goals and restrictions.

How well the resulting cost function reproduces the desired clinical priorities is determined by performing preliminary optimization runs and evaluating the tradeoffs chosen by the planner (as described below), or by experience. Often, the clinical tradeoffs which must be allowed, especially in the “middle priority” costlets, are not easy to determine at the start, and a number of trial optimization runs must be performed to discover what is achievable. During these trials, the physician or planner will alter one or several costlets and then assess the impact of the change on the overall plan which results. For example, the target overdose criterion (#3) can be relaxed to see if this allows a further decrease of dose to the spinal cord.

In practice, these techniques can be used in two general ways for any particular clinical protocol. The goal of our efforts to this point has been to include enough clinical information into the cost function so that once the cost function is determined for a particular protocol, all treatments for patients treated on that protocol can be planned using the same cost function. (A form of the cost function that can be used for a class of treatment protocols is referred to as a *class solution*.) At the other end of the spectrum, the physician/planner may also choose to derive a new cost function for each individual clinical case. Future research will study the potential benefit to individual patients by making individualized modifications to the class solution cost function.

3.4 Optimization algorithms

Recall that although the evaluators, and hence the costlets, are typically expressed in terms of doses to points, the doses can be computed as a known linear function of the beamlet weights. Therefore, once the treatment planner has constructed a cost function as described above, it serves as the means for comparing candidate treatment plans specified as the collection of beamlet weights within the automated planning system. With the cost function in place, an optimization algorithm is applied to minimize this function over all nonnegative beamlet weights. Depending on the nature of the evaluators and modifiers used in the construction of the cost function, various optimization methods can be used as a search mechanism to generate candidate plans. Our system allows the user to select an algorithm from a menu containing fast simulated annealing (Szu, 1987), dynamic hill climbing (Yuret, 1993, 1994), sequential quadratic programming (Gill, 2002), and quasi-Newton method (Byrd, 1995; Zhu, 1997).

4. Implementation and example

The optimization software described here has been designed and implemented as a prototype system to allow both clinical use and quantitative research in optimization for IMRT. The software has been written in C/C++ and runs on OpenVMS-based workstations. The user interface and graphical display features have been implemented using the Application Visualization System (Advanced Visual Systems Inc., Waltham, MA).

In this section we demonstrate the use and flexibility of the costlet-based optimization systems by developing a prototype cost function for patients with head-and-neck cancer. Most head-and-neck treatments involve several different target volumes with different priorities and numerous critical structures (Figure 1). The volumes were defined using images from X-ray CT and MR imaging. The primary planning target volume (PTV) was defined by geometrically expanding the macroscopic tumor volume (as visualized on the imaging studies) by 0.5 cm to account for patient positioning variations during the course of treatment. The left and right lymph node chains are at risk for disease, so they are also considered as targets but treated to a lower dose to reflect a lower level of risk. In addition to the spinal cord, it is important to “spare” the parotid glands which produce saliva (Eisbruch, 1998, 1999). Specific consideration of other clinically relevant structures, such as the oral cavity

Table 3. Clinical criteria for head-and-neck cancer treatment protocol.

Structure	Type	Min	Max	Mean
Planning target volume	<i>T</i> : Primary PTV	70 Gy	<77 Gy (1.1 × min)	<72 Gy (~1.03 × min)
Right boost volume PTV	<i>T</i> : High risk nodal volume	70 Gy	<77 Gy (1.1 × min)	<72 Gy (~1.03 × min)
Right nodal volume	<i>T</i> : Medium risk	64 Gy	<70.4 Gy (1.1 × min)	<66 Gy (~1.03 × min)
Left nodal volume	<i>T</i> : Low risk	60 Gy	<66 Gy (1.1 × min)	<62 Gy (~1.03 × min)
Right parotid gland	<i>N</i> : Healthy tissue			≤26 Gy
Left parotid gland	<i>N</i> : Healthy tissue			≤26 Gy
Spinal cord	<i>N</i> : Critical healthy tissue		45 Gy	

T = target, *N* = normal tissue.

and submandibular glands, were not performed here to keep the example straightforward and easy to present.

Table 3 contains a list of clinical goals set forth for the treatment in terms of the minimum, maximum, and/or mean doses to various structures. The relative priorities of meeting each of these goals are not specified, but will become apparent from the discussion below. Note that, as the treatment goals stated in the table suggest, the first four structures are volumes to be treated, while the remaining three are to be spared.

The beam arrangement used for this example was 7 equi-spaced coplanar beams placed axially around the patient spanning the entire 360 degrees (Figure 2(a)). The total number of beamlets was 1667 and the total number of points in all structures combined was 31,117.

In our initial example experiment, we attempted to meet the criteria for the four target volumes only, by designing a cost function which includes three costlets for each of the four treatment volumes. To penalize underdosing, *DVH_GE* evaluator with *Passthrough* modifier was used (100% of the volume should receive a dose greater than or equal to the minimum dose specified in the table). The *DVH_GE* evaluator is defined as follows: given a dose value d_1 and a parameter v such that it is desirable that at least $v\%$ of the volume of the structure under consideration receives a dose of d_1 or greater, we first find the dose value d_2 such that $v\%$ of the volume receives d_2 Gy or more in the current plan. Next, the value of the evaluator is computed as

$$E = \sum_{j:d_2 \leq d_j < d_1} (d_1 - d_j)^2,$$

where the summation is taken over all points j in the structure such that $d_2 \leq d_j < d_1$; the value of the evaluator is 0 if $v\%$ of the volume already receives a dose of d_1 Gy or higher (see Wu, 2000). A *DVH_LE* evaluator (defined similarly to *DVH_GE*, but computed as

$$E = \sum_{j:d_1 < d_j \leq d_2} (d_j - d_1)^2,$$

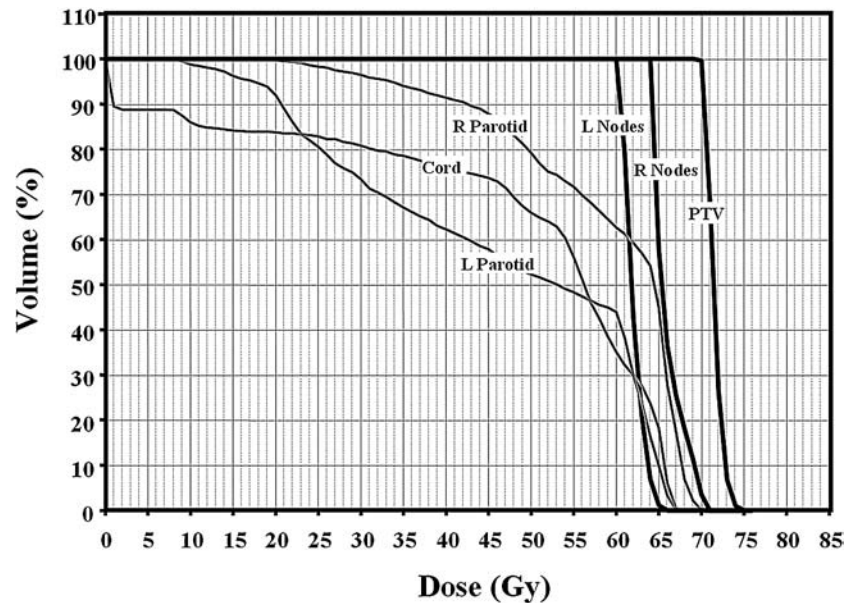


Figure 8. DVH results for optimization with costlets applied to target volumes only.

to penalize overdosing) with *Passthrough* modifier was used to penalize overdosing (0% of the volume should receive a dose less than or equal to the maximum dose). For the mean dose criteria, a *Mean* evaluator with a *CeilingPower* modifier (ceiling set at the desired maximum value of the mean and power 2) was used. A weighted sum of all the costlets (all with a weight of 1) was used to produce the overall cost function. An optimization performed using this cost function was able to find a plan that satisfied all target volume criteria (all costlet values were zero). The resulting dose-volume histograms are shown in Figure 8. The dose-volume histograms for the critical structures show significant overdosing, which was to be expected since clinical goals for those structures were not incorporated into the cost function.

To satisfy the maximum dose requirement for the spinal cord, *DVH_{LE}* evaluator with *Passthrough* modifier was added to the cost function, and the new optimal solution found. Both target and spinal cord criteria were met by the resulting plan. The resulting dose-volume histograms show a slight increase in heterogeneity of the dose received by the target volumes (Figure 9). The increase is a consequence of the beamlets overlapping the spinal cord being pushed to lower weights to meet the maximum dose criteria, and some of the remaining beamlets having a higher weight to compensate. The reduction in the overdosing of the parotids is also to be expected even though no costlets reflecting this criterion were explicitly included in the cost function, since weights of the beamlets that intersect both the spinal cord and the parotids are reduced.

Naturally, it is desirable to minimize the dose to all healthy structures. To demonstrate the impact of reduction of the dose to both the right and left parotid glands on the target volumes,

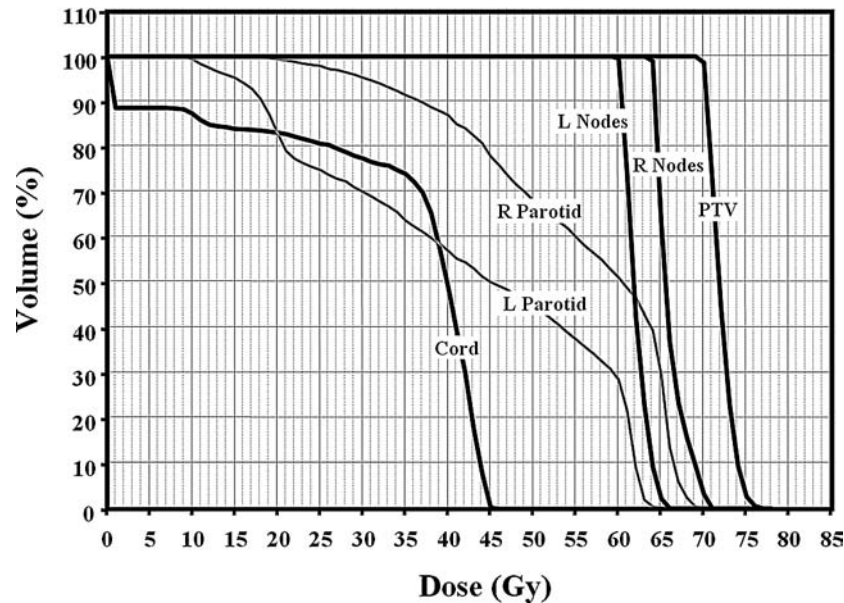


Figure 9. DVH results for optimization with costlets applied to target volumes and spinal cord.

costlets for these structures, consisting of *Mean* evaluator with Passthrough modifier, were added to the cost function with a weight of one. The *Mean* evaluator was used (instead of, for example, *Max*, or *DVH_LE*) since it was known that a portion of parotids overlaps with a target and would unavoidably receive some high dose. The resulting dose-volume histograms illustrate the tradeoff between target volume homogeneity and reduction of the dose to both parotids (Figure 10). Lowering the dose to the right parotid negatively affects the homogeneity of the right nodal target. The same is true for the left parotid and left nodal target, but to a lesser extent. This is illustrated by the solution obtained by removing the right parotid costlets from the cost function and re-optimizing the plan (Figure 11).

Taking into account the fact that the clinician would be willing to accept a treatment plan that delivers a mean dose of 26 Gy or less to a parotid, a *CeilingPower* modifier with a ceiling of 26 Gy and a power of 2 was applied to the *Mean* evaluator. Figure 12 shows that relaxing the penalty on the left parotid allows the goals for the left nodal target to be met for all but 1% of the volume.

When the above costlet is used for both left and right parotids, the amount of underdosing of the right nodal target becomes clinically unacceptable (Figure 13). To improve the homogeneity in the right nodal target, a *CeilingPower* modifier with a ceiling of 0 and successively increasing powers (2, 4, 10) was included (in addition to *DVH_GE* evaluator already present in the cost function). Figure 14 illustrates the results of these three optimizations and shows that only small improvements in the homogeneity of the right nodal target are gained for modest increases in the mean dose to the right parotid. This is due to the degree of geometric overlap between these two structures and the differences in their

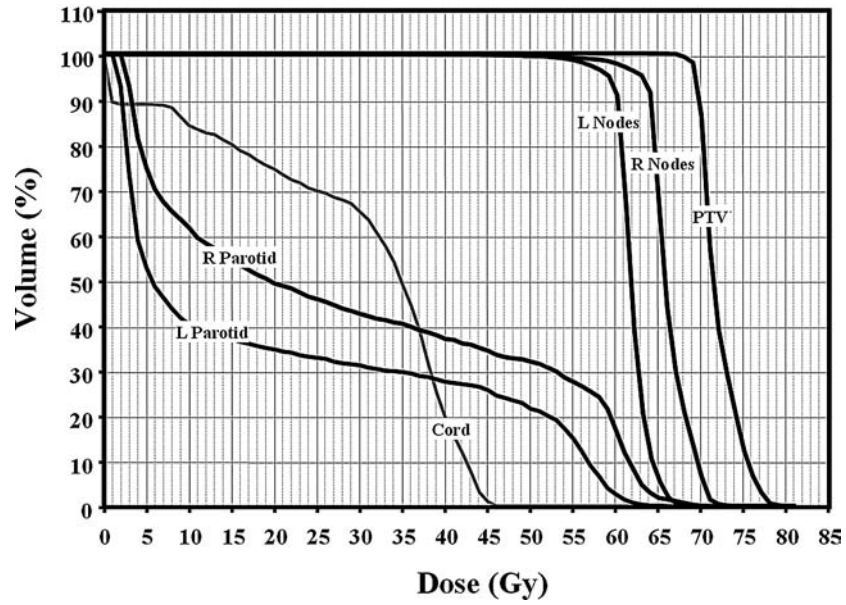


Figure 10. DVH results for optimization with costlets applied to all structures.

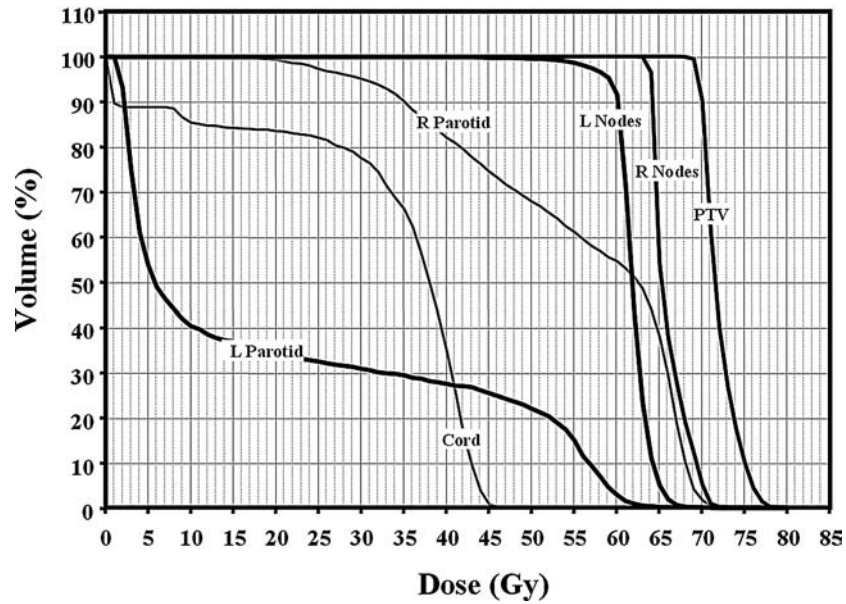


Figure 11. DVH results for optimization with costlets applied to all structures except right parotid.

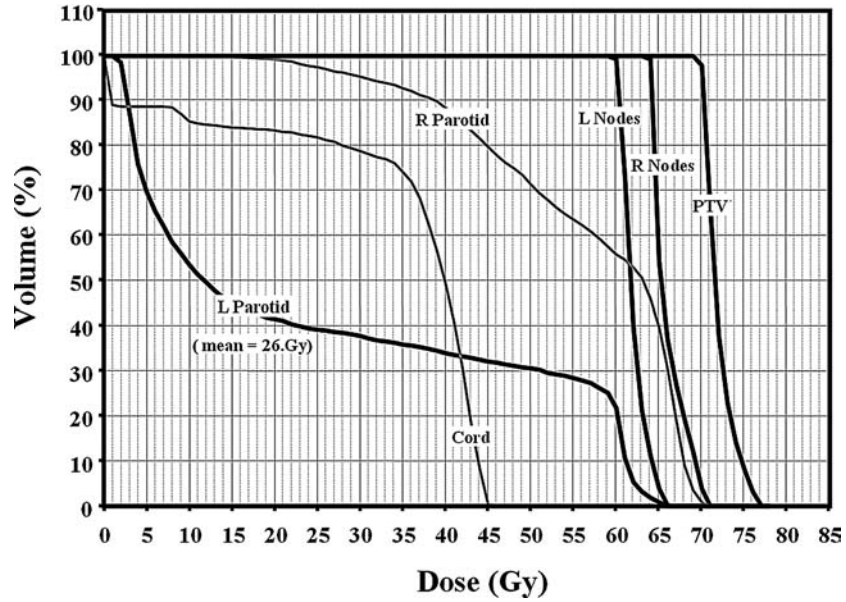


Figure 12. DVH results for optimization with clinically acceptable mean dose for left parotid.

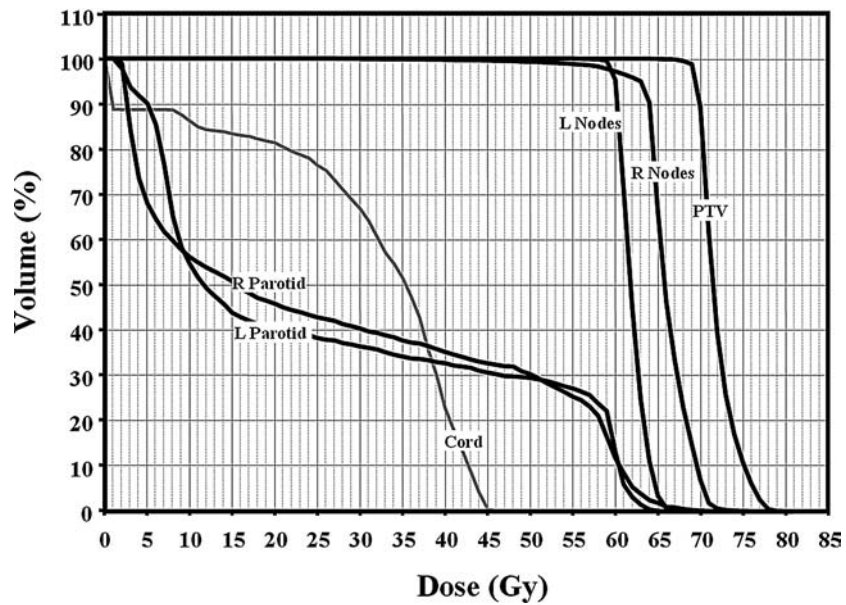


Figure 13. DVH results for optimization with clinically acceptable mean for left and right parotid.

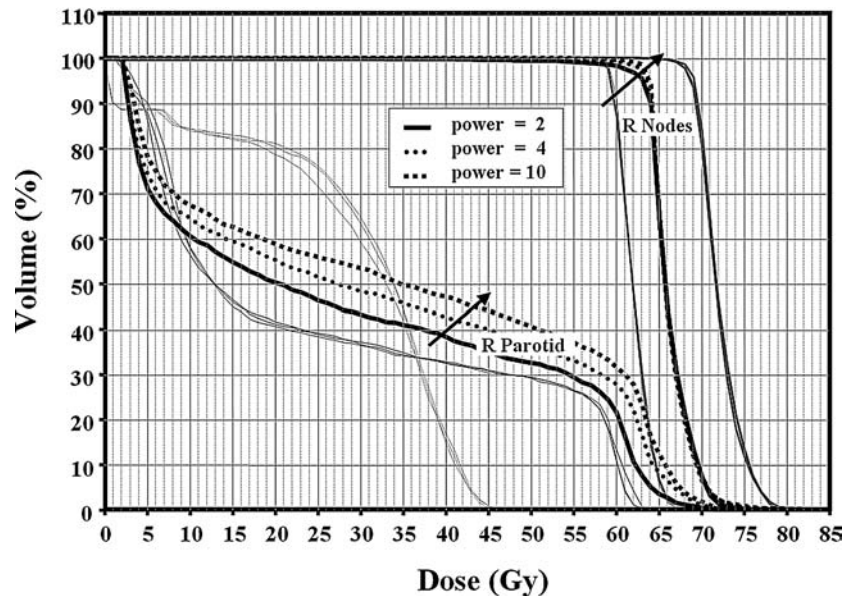


Figure 14. Comparison of effects of different modifier powers applied to right nodes.

volume relative to this overlap. The ability to interactively specify different cost functions, and find and analyze resulting solutions, helps the clinician to understand the advantages and disadvantages of different trade-offs, and to elucidate the underlying reasons for these trade-offs.

As a final example of the use and flexibility of the optimization system, the costlets that were introduced in an attempt to achieve a mean dose of at most 26 Gy for both parotids by using *Mean* evaluator in the example above were replaced with an *NTCP* evaluator for the right parotid. A *CeilingPower* modifier was used to produce a dose distribution that still met the mean dose criteria. The dose-volume histograms comparing these two solutions are shown in Figure 15. Although the trade-offs with the other structures are very similar, the shapes of the right parotid dose-volume histograms are very different. This difference is due to the sigmoid shape of the *NTCP* evaluator, which penalizes higher doses much more significantly than lower doses (Lyman, 1987).

5. Discussion and conclusions

Although work on automated optimization for radiotherapy planning has been pursued for 35 years, dramatic progress has been made mostly over the past 10 years. One aspect of this progress has been development of clinically relevant cost functions, since the cost function is the key which drives an inverse plan toward the “optimal” result. A cost function which accurately reflects clinical goals is a fundamental and crucial part of all plan optimization efforts. Various cost functions have been reported, including linear (Rosen, 1991;

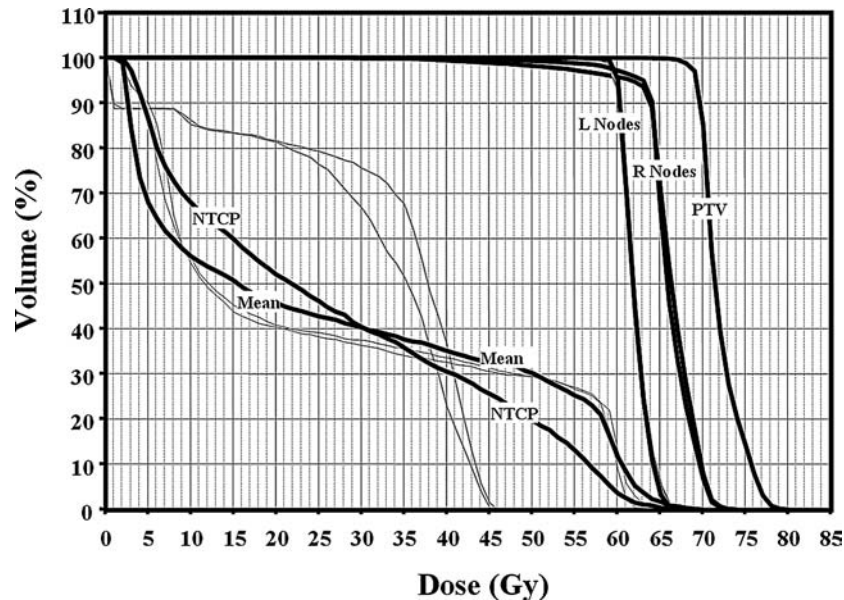


Figure 15. Comparison of effects of mean versus NTCP costlets for right parotid.

Langer, 1990) or quadratic (Holmes, 1994; De Wagter, 1998; Oldham, 1995; Redpath, 1976) functions, the probability of uncomplicated control (P^+) (Kallman, 1992; Agren, 1990; Soderstrom, 1993), and other methods based on biological indices (such as TCP, NTCP, V_{eff}).

Each type of cost function has its inherent advantages and disadvantages, depending on the optimization system, clinical goals, clinical knowledge, expectations, and other features. Dose-based cost functions are limited in that they do not take into account biological response (Wang, 1995), however, they also do not have the large uncertainties which are currently associated with biologically-based functions (Webb, 1997). In addition, they may be simpler and faster to calculate than some other candidate cost functions (Wang, 1995). Quadratic least-squares dose-based functions (Holmes, 1994) have the advantage that they can always achieve some solution (albeit sometimes a poor solution), and are implemented efficiently into some search strategies, while at the same time they have a fundamental difficulty achieving rigorous dosimetric limits for particular structures (Mohan, 1994).

The various limitations and advantages of the many published methods for optimization cost functions have led us to develop the generalized cost function methodology described in this paper. This method allows the planner to create a cost function by selecting appropriate components from a wide variety of techniques. Individual "costlets" are used to describe individual aspects of the clinical goals of the planning. The planner is given the freedom to combine different types of costlets, to use different methods for combining costlets, and to define rigid constraints as well as costs. In addition to use for clinical IMRT planning, the methodology has been designed to allow investigation of various parts of the IMRT planning

process, including possible improvements in cost function formulation and optimization methods (Vineberg, 1999, 2000, 2002) as well as comparing optimization-based planning of IMRT and non-IMRT conformal therapy (Fraass, 2000).

It is expected that automated optimization of IMRT will really begin to make a significant impact as clinical knowledge and biological effects are quantitatively incorporated into the planning process. There are still many questions that remain concerning how this biological information should be incorporated into the optimization system. The most straightforward method might involve the use of the probability of uncomplicated control implemented using the P^+ function. However, there are a number of limitations to this simple function, including our incomplete knowledge of the details of how each structure's NTCP varies with the parameters of different plan techniques and how to weight the various complication endpoints for each structure. In addition, the data used to create the appropriate TCP functions for each grade of tumor is also extremely limited. Third, it is not clear that using P^+ as a cost function for optimization is the most efficient way to obtain the optimal result, given that the sensitivity of the various biological models may not be large enough to force the optimization toward the desired result. It is also possible that control and/or complication endpoints might be correlated (i.e., a late complication like lung fibrosis might be correlated with a more immediate complication like pneumonitis), and this should be taken into account in probabilities and weighting of the individual costlets. Finally, there are additional costlets which are probably important in the overall optimization of a plan, including issues like treatment delivery time and complexity, the stability of the plan against various patient setup, organ motion, and treatment delivery uncertainties. The approach chosen here, in contrast to the P^+ approach, allows any kind of biological or clinical information to be incorporated into the problem. In addition, this approach allows use of limited data without forcing the entire cost function to be driven by biological data: it is possible to incorporate the use of a single organ NTCP into the problem, while other pieces of the cost function remain dose or dose-volume based.

One of the aims of computer-assisted optimization is to help the planner and physician develop better treatment plans for patient treatments. In the context of this paper on cost functions and optimization, it is interesting to note that almost all planning protocols in use today, including virtually all formal planning studies, consist almost entirely of constraints, not optimization criteria. Constraint-defined protocols have been the basis of many of the clinical study results which have helped advance the field of radiation oncology, and they should not be removed without careful analysis of our ability to continue to perform quantitative analysis of the results of planning and treatment-related studies which are performed based on non-constraint defined treatments. However, these constraint-based protocol designs make the introduction of formally-optimized planning into the planning process a difficult task.

Certainly, inverse planning performed by systems which use simple quadratic dose difference cost functions often leads to dose distributions which do not satisfy the dose uniformity constraints typically required for standard conformal therapy (dose uniformity inside the PTV of $\pm 5\%$). There are many examples of "optimized" treatments generated with an inverse planning system which contain dose cold spots (under-dosed regions in the target(s)) as large as 30 or 40% and dose hot spots (unwanted over-dosed regions) of 30–40% inside

the PTV (Boyer, 1997; Butler, 1999; Chao, 2000). Often, these large dose variations are considered to be a result of the optimization process. However, these differences are actually due to the specific construction of the cost function which was used, and they can be avoided if desired (Vineberg, 2000). The examination of the different tradeoffs which result from using hard optimization constraints (with more quantitatively-controlled results) versus using simple costlets (possibly leading to more easily achieved solutions, but with less precise adherence to the desired limitations) must be performed, so we can better understand when the two different methods should be used. When rigid constraints are necessary, and when more flexibility can be tolerated, is an important clinical issue for each protocol. It is critical that future conformal therapy trials which are based on inverse planning and/or optimized IMRT plans correctly handle the potentially qualitative differences in normal tissue and tumor dose distributions which result from the different kinds of cost functions which may be used.

One should also note that general optimization of most planning problems will require numerous costlets and/or constraints which are now only implicitly included in planning or study protocols. Due to the complexity of the overall optimization problem, many current optimization approaches make use of a very limited degree of flexibility in the possible solutions, and these "arbitrary" decisions may seriously affect the overall quality of the "optimized" treatment plan which results from the optimization process. Further, many basic assumptions of treatment planning which are known intuitively by the experienced human treatment planner must be explicitly integrated into the overall optimization cost function. These issues may include full knowledge of machine-table-patient combinations which are possible, delivery time limits, safety limits, and complexity limits, among others. The implementation of more clinically-intuitive interactions with the optimization strategy may also help improve the quality and efficiency of the planning process and results. These more sophisticated developments will become clearer as clinical studies begin to give feedback on which to base costlet and cost function definition.

Acknowledgments

Supported in part by NIH P01-CA59827 and NSF CCF-0306240.

References

- A. Agren, A. Brahme, and I. Turesson, "Optimization of uncomplicated control for head and neck tumors," *International Journal of Radiation Oncology Biology Physics* vol. 19, pp. 1077–1085, 1990.
- M. Alber and F. Nusslin, "Optimization of intensity modulated radiotherapy under constraints for static and dynamic MLC delivery," *Physics in Medicine and Biology* vol. 46, pp. 3229–39, 2001.
- M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 2nd ed (Wiley, 1993).
- T. Bortfeld, A. L. Boyer, W. Schlegel, D. L. Kahler, and T. J. Waldron, "Realization and verification of 3-dimensional conformal radiotherapy with modulated fields," *International Journal of Radiation Oncology Biology Physics* vol. 30, pp. 899–908, 1994.
- T. Bortfeld and W. Schlegel, "Optimization of beam orientations in radiation-therapy—some theoretical considerations," *Physics in Medicine and Biology* vol. 38, pp. 291–304, 1993.

- A. L. Boyer, P. Geis, W. Grant, and M. Carol, "Modulated beam conformal therapy for head and neck tumors," *International Journal of Radiation Oncology Biology Physics* vol. 39, pp. 227–236, 1997.
- A. Brahme, "Design principles and clinical possibilities with a new generation of radiation therapy equipment. A review," *Acta Oncologica*, vol. 26, pp. 403–412, 1987.
- A. Brahme, "Optimization of radiation therapy and the development of multileaf collimation," *International Journal of Radiation Oncology Biology Physics* vol. 25, pp. 373–375, 1993.
- A. Brahme, "Optimization of Stationary and Moving Beam Radiation-Therapy Techniques," *Radiotherapy and Oncology* vol. 12, pp. 129–140, 1988.
- A. Brahme and A. K. Agren, "Optimal Dose Distribution for Eradication of Heterogeneous Tumors," *Acta Oncologica* vol. 26, pp. 377–385, 1987.
- C. Burman, C. S. Chui, G. Kutcher, S. Leibel, M. Zelefsky, T. LoSasso, S. Spirou, Q. Wu, J. Yang, J. Stein, R. Mohan, Z. Fuks, and C. C. Ling, "Planning, delivery, and quality assurance of intensity-modulated radiotherapy using dynamic multileaf collimator: A strategy for large-scale implementation for the treatment of carcinoma of the prostate," *International Journal of Radiation Oncology, Biology, Physics* vol. 39, pp. 863–873, 1997.
- E. B. Butler, B. S. Teh, W. H. Grant, B. M. Uhl, R. B. Kuppersmith, J. K. Chiu, D. T. Donovan, and S. Y. Woo, "Smart boost: A new accelerated fractionation schedule for the treatment of head and neck cancer with intensity modulated radiotherapy," *International Journal of Radiation Oncology Biology Physics* vol. 45, pp. 21–32, 1999.
- R. H. Byrd, P. Lu, and J. Nocedal, "A limited memory algorithm for bound constrained optimization," *SIAM Journal on Scientific and Statistical Computing* vol. 6, pp. 1190–1208, 1995.
- M. Carol, "Integrated 3-D conformal multivane intensity modulation delivery system for radiotherapy," in *Proceedings of the XIth International Conference on the Use of Computers in Radiation Therapy*, A. Hounsell, J. Wilkinson, and P. Williams, Eds. Manchester, UK: Medical Physics, pp. 172, 1994.
- M. P. Carol, "PeacockTM: A system for planning and rotational delivery of intensity-modulated fields," *International Journal of Imaging Systems and Technology* vol. 6, pp. 56–61, 1995.
- K. S. C. Chao, D. Low, C. A. Perez, and J. A. Purdy, "Intensity-modulated radiation therapy in head and neck cancer: The Mallinckrodt experience," *Int. J. Cancer (Rad. Onc. Invest.)* vol. 90, pp. 92–103, 2000.
- E. M. Damen, M. J. Brugmans, A. van der Horst, L. Bos, J. V. Lebesque, B. J. Mijnheer, D. L. McShan DL, B. A. Fraass, and M. L. Kessler, "Planning, computer optimization, and dosimetric verification of a segmented irradiation technique for prostate cancer," *International Journal of Radiation Oncology Biology Physics* vol. 49, pp. 1183–95, 2001.
- C. De Wagter, C. O. Colle, L. G. Fortan, B. B. Van Duyse, D. L. Van den Berge, and W. J. De Neve, "3D conformal intensity-modulated radiotherapy planning: Interactive optimization by constrained matrix inversion," *Radiotherapy and Oncology* vol. 47, pp. 69–76, 1998.
- A. Eisbruch, L. H. Marsh, M. K. Martel, J. A. Ship, R. Ten Haken, A. T. Pu, B. A. Fraass, and A. S. Lichter, "Comprehensive irradiation of head and neck cancer using conformal multisegmental fields: Assessment of target coverage and noninvolved tissue sparing," *International Journal of Radiation Oncology Biology Physics* vol. 41, pp. 559–568, 1998.
- A. Eisbruch, R. K. Ten Haken, H. M. Kim, L. H. Marsh, and J. A. Ship, "Dose, volume, and function relationships in parotid salivary glands following conformal and intensity-modulated irradiation of head and neck cancer," *International Journal of Radiation Oncology Biology Physics* vol. 45, pp. 577–587, 1999.
- M. C. Ferris, R. R. Meyer, and W. D'Souza, "Radiation treatment planning: Mixed integer programming formulations and approaches," Computer Sciences Department, University of Wisconsin, Madison, Optimization Technical Report 02–08, 2002.
- B. A. Fraass and D. L. McShan, "3-D treatment planning: I Overview of a clinical planning system," in *The Use of Computers in Radiation Therapy*, I. A. D. Bruinvis, F. H. van_der_Giessen, H. J. van_Kleffens, and F. W. Wittkamper, Eds. North-Holland: Elsevier Science Publishers BV, 1987, pp. 273–276, 1987a.
- B. A. Fraass, D. L. McShan, R. K. TenHaken, and K. M. Hutchins, "3-D treatment planning: V. A Fast 3-D photon calculation model," in *The Use of Computers in Radiation Therapy*, I. A. D. Bruinvis, F. H. van_der_Giessen, H. J. van_Kleffens, and F. W. Wittkamper, Eds. North-Holland: Elsevier Science Publishers BV, pp. 521–525, 1987b.
- B. A. Fraass, D. L. McShan, and M. L. Kessler, "Dose-Based Conformal Field Shaping Using Automated Optimization," in *Proceedings of the XIIIth International Conference On the Use of*

- Computers In Radiotherapy*, Heidelberg, Germany: T. Bortfeld and W. Schlegel, Eds., pp. 32–35, 2000.
- B. A. Fraass, D. L. McShan, and K. J. Weeks, “3-D treatment planning: III. Complete Beam’s-Eye-View planning capabilities,” in *The Use of Computers in Radiation Therapy*, I. A. D. Bruinvis, F. H. van_der_Giessen, H. J. van_Kleffens, and F. W. Wittkamper, Eds., North-Holland: Elsevier Science Publishers BV, pp. 193–196, 1987.
- P. E. Gill, W. Murray, and M. A. Saunders, “SNOPT: An SQP algorithm for large-scale constrained optimization,” *SIAM Journal on Optimization* vol. 12, pp. 979–1006, 2002.
- M. Goitein, et al., “Evaluation of treatment planning for particle beam radiotherapy: report of the working groups,” Radiotherapy Development Branch, Radiation Research Program, Division of Cancer Treatment, National Cancer Institute, Bethesda MD, 1987.
- D. Greene and P. C. Williams, *Linear Accelerators for Radiation Therapy, Series in Medical Physics and Biomedical Engineering*, pp. 194, 1985.
- H. W. Hamacher and K. H. Kufer, “Inverse radiation therapy planning—a multiple objective optimization approach,” *Discrete Applied Mathematics* vol. 118, pp. 145–161, 2002.
- A. Holder, “Designing radiotherapy plans with elastic constraints and interior point methods,” *Health Care and Management Science* vol. 6, pp. 5–16, 2003.
- A. Holder, “Partitioning multiple objective optimal solutions with applications in radiotherapy design,” Trinity University, San Antonio, Mathematics Technical Report 54, 2001.
- T. Holmes, T. Mackie, D. Simpkin, and P. Reckwerdt, “A unified approach to the optimization of brachytherapy and external beam therapy,” *International Journal of Radiation Oncology, Biology, Physics* vol. 20, pp. 859–873, 1991.
- T. Holmes and T. R. Mackie, “A filtered backprojection dose calculation method for inverse planning,” *Medical Physics* vol. 21, pp. 303–313, 1994.
- T. W. Holmes, T. R. Mackie, and P. Reckwerdt, “An iterative filtered backprojection inverse treatment planning algorithm for tomotherapy,” *International Journal of Radiation Oncology, Biology, Physics* vol. 32, pp. 1215–1225, 1995.
- ICRU Report 50. Prescribing, Recording and Reporting Photon Beam Therapy, International Commission on Radiation Units and Measurements, 7910 Woodmont Avenue, Bethesda, Maryland 20814, USA, 1993.
- P. Kallman, B. Lind, and A. Brahme, “An algorithm for maximizing the probability of complication-free tumour control on radiation therapy,” *PMB* vol. 37, pp. 871–890, 1992.
- M. L. Kessler, S. Pitluck, P. Petti, and J. R. Castro, “Integration of Multimodality Imaging Data for Radiotherapy Treatment Planning,” *International Journal of Radiation Oncology Biology Physics* vol. 21, pp. 1653–1667, 1991.
- E. A. Krueger, B. A. Fraass, D. L. McShan, R. Marsh, and L. J. Pierce, “Potential gains for irradiation of chest wall and regional nodes with intensity modulated radiotherapy,” *International Journal of Radiation Oncology Biology Physics* vol. 56, pp. 1023–37, 2003.
- G. J. Kutcher and C. Burman, “Calculation of Complication Probability Factors for Non-Uniform Normal Tissue Irradiation—the Effective Volume Method,” *International Journal of Radiation Oncology Biology Physics* vol. 16, pp. 1623–1630, 1989.
- K. H. Küfer, A. Scherrer, M. Monz, F. Alonso, H. Trinkaus, T. Bortfeld, and C. Thieke, “Intensity-modulated radiotherapy—a large scale multi-criteria programming problem,” *OR Spectrum* vol. 25, pp. 223–249, 2003.
- M. Langer and J. Leong, “Optimization of Beam Weights under Dose-Volume Restrictions,” *International Journal of Radiation Oncology Biology Physics* vol. 13, pp. 1255–1260, 1987.
- M. Langer, S. S. Morrill, and R. Lane, “A test of the claim that plan rankings are determined by relative complication and tumor-control probabilities,” *International Journal of Radiation Oncology Biology Physics* vol. 41, pp. 451–457, 1998.
- E. K. Lee, T. Fox, and I. Crocker, “Integer programming applied to intensity-modulated radiation therapy treatment planning,” *Annals of Operations Research* vol. 119, pp. 165–181, 2003.
- J. H. Lim, M. C. Ferris, S. J. Wright, D. M. Shepard, and M. A. Earl, “An optimization framework for conformal radiation treatment planning,” Computer Sciences Department, University of Wisconsin, Madison, Optimization Technical Report 0210, 2002.
- J. T. Lyman, “Complication probability as assessed from dose volume histograms,” *Rad Res* vol. 104, pp. 5–13, 1985.

- J. T. Lyman, "Normal tissue complication probabilities—variable dose per fraction," *International Journal of Radiation Oncology Biology Physics* vol. 22, pp. 247–250, 1992.
- J. T. Lyman and A. B. Wolbarst, "Optimization of radiation therapy III: A method of assessing complication probabilities from dose volume histograms," *International Journal of Radiation Oncology Biology Physics* vol. 13, pp. 103–109, 1987.
- T. R. Mackie, J. W. Scrimger, and J. J. Battista, "A convolution method of calculating dose for 15-MV X-rays," *Medical Physics* vol. 12, pp. 188–196, 1985.
- D. L. McShan and B. A. Fraass, "3-D treatment planning: II. Integration of gray scale images and solid surface graphics," in *The Use of Computers in Radiation Therapy*, I. A. D. Bruinvis, F. H. van_der_Giessen, H. J. van_Kleffens, and F. W. Wittkamper, Eds. North-Holland: Elsevier Science Publishers BV, pp. 41–44, 1987.
- D. L. McShan, B. A. Fraass, and A. S. Lichter, "Full integration of the beam's eye view concept into computerized treatment planning," *International Journal of Radiation Oncology Biology Physics* vol. 18, pp. 1485–1494, 1990.
- G Meedt, M Alber, and F Nusslin, "Non-coplanar beam direction optimization for intensity-modulated radiotherapy," *Physics in Medicine and Biology* vol. 48, pp. 2999–3019, 2003.
- C. R. Meyer, J. L. Boes, B. Kim, P. H. Bland, et al., "Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin plate spline warped geometric deformations," *Medical Image Analysis* vol. 1, no. 3, pp. 195–206, 1997.
- D. Michalski, Y. Xiao, Y. Censor, and J. M. Galvin, "The dose-volume constraint satisfaction problem for inverse treatment planning with simultaneous subgradient projection method," *Medical Physics* vol. 30, pp. 1334–1334, 2003.
- R. Mohan, G. S. Mageras, B. Baldwin, L. J. Brewster, G. J. Kutcher, S. Leibel, C. M. Burman, C. C. Ling, and Z. Fuks, "Clinically relevant optimization of 3-D conformal treatments," *Medical Physics* vol. 19, pp. 933–944, 1992.
- R. Mohan, X. Wang, and A. Jackson, "Optimization of 3-D conformal radiation treatment plans," *Frontiers of Radiation Therapy & Oncology* vol. 29, pp. 86–103, 1996.
- R. Mohan, X. H. Wang, A. Jackson, T. Bortfeld, A. L. Boyer, G. J. Kutcher, S. A. Leibel, Z. Fuks, and C. C. Ling, "The potential and limitations of the Inverse Radiotherapy Technique," *Radiotherapy and Oncology* vol. 32, pp. 232–248, 1994.
- A. Niemierko, "Reporting and analyzing dose distributions: The concept of equivalent uniform dose," *Medical Physics* vol. 24, pp. 103–110, 1997.
- A. Niemierko and M. Goitein, "Implementation of a model for estimating tumor control probability for an inhomogeneously irradiated tumor," *Radiotherapy and Oncology* vol. 29, pp. 140–147, 1993.
- M. Oldham, A. Neal, and S. Webb, "A comparison of conventional forward planning with inverse planning for 3-D conformal radiotherapy of the prostate (Vol 35, Pg 248, 1995)," *Radiotherapy and Oncology* vol. 37, pp. 171–172, 1995.
- F. Preciado-Walters, M. Langer, R. Rardin, and V. Thai, "A coupled column generation, mixedinteger approach to optimal planning of intensity modulated radiation therapy for cancer," *Mathematical Programming*, 2004.
- W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, "Numerical Recipes", 2nd ed., Cambridge University Press, Cambridge, 1992.
- A. T. Redpath, B. L. Vickery, and D. H. Wright, "New technique for radiotherapy planning using Quadratic programming," *Physics in Medicine and Biology* vol. 21, pp. 781–791, 1976.
- H. E. Romeijn, R. K. Ahuja, J. F. Dempsey, A. Kumar, and J. G. Li, "A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planning," *Physics in Medicine and Biology* vol. 48, pp. 3521–3542, 2003.
- C. G. Rowbottom, C. M., Nutting, and S. Webb., "Beam-orientation optimization of intensity-modulated radiotherapy: Clinical application to parotid gland tumours," *Radiotherapy and Oncology* vol. 59, pp. 169–77, 2001.
- Y. Seppenwoolde, M. Engelsman, K. De Jaeger, S. H. Muller, P. Baas, J. S. A. Belderbos, L. J. Boersma, D. L. McShan, B. A. Fraass, M. L. Kessler, and J. V. Lebesque, "Optimizing radiation treatment plans for lung cancer using functional information," *Radiotherapy and Oncology* vol. 63, pp. 165–77, 2002.
- D. M. Shepard, M. C. Ferris, G. H. Olivera, and T. R. Mackie, "Optimizing the delivery of radiation therapy to cancer patients," *Siam Review* vol. 41, pp. 721–744, 1999.

- S. Sodertrom and A. Brahme, "Optimization of the Dose Delivery in a Few Field Techniques using Radiobiological Objective Functions," *Medical Physics* vol. 20, pp. 1201–1210, 1993.
- S. V. Spirou and C. S. Chui, "A gradient inverse planning algorithm with dose-volume constraints," *Medical Physics* vol. 25, pp. 321–333, 1998.
- H. Szu and R. Hartley, "Fast simulated annealing," *Physics Letters A* vol. 122, pp. 157–162, 1987.
- R. K. Tenhaken, A. F. Thornton, H. M. Sandler, M. L. Lavigne, D. J. Quint, B. A. Fraass, M. L. Kessler, and D. L. Meshan, "A quantitative assessment of the addition of MRI to CT-based, 3-D treatment planning of Brain-Tumors," *Radiotherapy and Oncology* vol. 25, pp. 121–133, 1992.
- K. Vineberg, B. A. Fraass, M. L. Kessler, D. L. McShan, and A. Eisbruch, "Parotid sparing without sacrificing target dose uniformity using optimized beamlet IMRT," presented at American Society of Therapeutic Radiology and Oncology, Boston MA., 2000.
- K. A. Vineberg, A. Eisbruch, M. M. Coselmon, D. L. McShan, M. L. Kessler, and B. A. Fraass, "Is uniform target dose possible in IMRT plans in the head and neck?," *International Journal of Radiation Oncology Biology Physics* vol. 52, pp. 1159–1172, 2002.
- K. A. Vineberg, M. K. Martel, M. L. Kessler, D. L. McShan, J. J. Kim, H. S. Sandler, and B. A. Fraass, "Dose Escalation of Brain Tumors to 100+ Gy using automated IMRT optimization," *International Journal of Radiation Oncology Biology Physics* vol. 45, pp. 270, 1999.
- X. H. Wang, R. Mohan, A. Jackson, S. A. Leibel, Z. Fuks, and C. C. Ling, "Optimization of intensity-modulated 3D conformal treatment plans based on biological indices," *Radiotherapy & Oncology* vol. 37, pp. 140–152, 1995.
- S. Webb, "Optimisation of conformal radiation therapy dose distributions by simulated annealing," *Physics in Medicine and Biology* vol. 34, pp. 1349–1370, 1989.
- S. Webb, "Optimization by simulated annealing of 3-Dimensional conformal treatment planning for radiation-fields defined by a Multileaf Collimator," *Physics in Medicine and Biology* vol. 36, pp. 1201–1226, 1991.
- S. Webb, *The Physics of Conformal Radiotherapy, Advances in Technology*. Bristol: IOP Publisher Ltd, 1997.
- Q. Wu, R. Mohan, A. Niemierko A, and R. Schmidt-Ullrich, "Optimization of intensity-modulated radiotherapy plans based on the equivalent uniform dose," *International Journal of Radiation Oncology Biology Physics* vol. 52, pp. 224–35, 2002.
- Q. Wu and R. Mohan, "Algorithms and functionality of an intensity modulated radiotherapy optimization system," *Medical Physics* vol. 27, pp. 701–11, 2000.
- Y. Xiao, D. Michalski, J. M. Galvin, and Y. Censor, "The least-intensity feasible solution for aperture-based inverse planning in radiation therapy," *Annals of Operations Research* vol. 119, pp. 183–203, 2003.
- L. Xing L, J.G. Li, S. Donaldson S, Q. T. Le, and A. L. Boyer, "Optimization of importance factors in inverse planning," *Physics in Medicine and Biology* vol. 44, pp. 2525–36, 1999.
- D. Yuret, "From genetic algorithms to efficient optimization," Technical Report 1569, MIT AI Laboratory, 1994.
- D. Yuret and M. de la Maza, "Dynamic hill climbing: Overcoming the limitations of optimization techniques," In Second Turkish Symposium on Artificial Intelligence and Artificial Neural Networks, pp. 208–212, 1993.
- C. Zhu, R. H. Byrd, and J. Nocedal, "L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization," *ACM Transactions on Mathematical Software* vol. 23, pp. 550–560, 1997.