A characterization of robust radiation therapy treatment planning methods—from expected value to worst case optimization

Albin Fredriksson[a)]
Optimization and Systems Theory, Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden and RaySearch Laboratories, Sveavägen 25, SE-111 34 Stockholm, Sweden

(Received 29 February 2012; revised 27 June 2012; accepted for publication 29 June 2012; published 31 July 2012)

Purpose: To characterize a class of optimization formulations used to handle systematic and random errors in radiation therapy, and to study the differences between the methods within this class.
Methods: The class of robust methods that can be formulated as minimax stochastic programs is studied. This class generalizes many previously used methods, ranging between optimization of the expected and the worst case objective value. The robust methods are used to plan intensity-modulated proton therapy (IMPT) treatments for a case subject to systematic setup and range errors, random setup errors with and without uncertain probability distribution, and combinations thereof. As reference, plans resulting from a conventional method that uses a margin to account for errors are shown.
Results: For all types of errors, target coverage robustness increased with the conservativeness of the method. For systematic errors, best case organ at risk (OAR) doses increased and worst case doses decreased with the conservativeness. Accounting for random errors of fixed probability distribution resulted in heterogeneous dose. The heterogeneities were reduced when uncertainty in the probability distribution was accounted for. Doing so, the OAR doses decreased with the conservativeness. All robust methods studied resulted in more robust target coverage and lower OAR doses than the conventional method.
Conclusions: Accounting for uncertainties is essential to ensure plan quality in complex radiation therapy such as IMPT. The utilization of more information than conventional in the optimization can lead to robust target coverage and low OAR doses. Increased target coverage robustness can be achieved by more conservative methods. [http://dx.doi.org/10.1118/1.4737113]

Key words: robust planning, uncertainty, optimization, IMPT

I. INTRODUCTION

A key aspect of external beam radiation therapy is the collocation of the patient anatomy and the treatment beams. If potential maladjustments are not accounted for in the planning process, the delivered treatment may degrade severely compared to the planned. In conventional radiation therapy, region of interest (ROI) margins are used to account for errors. However, the effectiveness of margins decreases as the complexity of the treatment increases. Other means of handling the uncertainties are therefore often preferable for modulated treatments. Intensity-modulated radiation therapy (IMRT) is one modality allowing for high modulation and thereby complex treatments. Active scanning of ion beams—such as intensity-modulated proton therapy (IMPT)—allows for even more modulation and is thus highly susceptible to errors.[1,2]

Different methods using additional information in the optimization have been proposed to overcome the shortcomings of margins.[3–5] Not surprisingly, such methods appear to enable better plans in general. Still, the resulting differences between the methods used for robustness remain principally unknown. In this paper, the aim is to study the effects of a class of robust methods for systematic and random errors, encompassing many of those presented in the literature. The methods studied are those that can be formulated as minimax stochastic optimization problems.[6,7] These range between expected value and worst case optimization.

Many methods for robust radiation therapy are based on expected value optimization.[3,4] For physical treatment planning criteria, such as quadratic dose deviations, optimization of the expected value with respect to systematic, and random errors was performed by Unkelbach and Oelfke.[4] In a previous paper,[5] we too used physical criteria, but minimized the worst case penalty in a set of systematic error scenarios. Chen *et al.*[8] included worst case and expected value optimization of physical functions with respect to systematic errors in a linear programming multicriteria optimization framework, thus enabling interactive navigation between the optimum with respect to robust optimization functions and the optimum with respect to nominal functions.

In this paper, the minimax stochastic methods are used with physical criteria that penalize dose deviations quadratically. To get a clear view of the different methods, their resulting dose distributions on a two-dimensional phantom treated with IMPT are studied. It should, however, be noted that the methods are applicable to any modality. First, the methods are used to account for systematic and random errors individually. In the case of random errors, the importance of taking uncertainty in the probability distribution into account is highlighted, and the minimax stochastic methods are applied

to do so. Putting things together, systematic errors and random errors with uncertain standard deviation are simultaneously handled in the optimization, and the resulting plans are studied.

Robust methods not belonging to the class considered here have also been used previously. For instance, robust linear programming[9,10] has been used to achieve robust IMRT plans: Chu *et al.*[11] accounted for organ motion of uncertain distribution to achieve high probability of sufficient dose to each target voxel considered individually. Chan *et al.*[12] ensured that all target voxels received sufficient expected dose for all probability distributions within a given set. Accounting for multiple probability distributions is similar to the formulation for random errors considered here. A difference lies in that the method in this paper operates on general nonlinear penalty functions and not directly on the voxel doses.

II. METHODS

II.A. Uncertainties

Systematic errors and random errors are considered. The errors may originate from any source, such as setup uncertainty, range uncertainty, or organ motion. The systematic (treatment preparation) errors are the same in each treatment fraction, whereas the random (treatment execution) errors are randomly realized anew in each fraction. Intrafractional errors are neglected. The random errors are assumed to be independent and identically distributed (i.i.d.), but of uncertain distribution. It is assumed that the uncertain distributions can be characterized by their standard deviations, which are random variables.

The possible realizations of the uncertainties are discretized into scenarios. The notation for the scenario index sets, the random variables picking scenarios from these sets, and the probability distributions of the random variables for the different types of uncertainties is specified in Table I. *A priori* probability distributions are denoted by Latin letters and uncertain probability distributions, which are variables in the optimizations, are denoted by Greek letters. The random error probability distribution q is conditioned on the standard deviation σ. When there are n fractions, the set $\mathcal{T}$ of random error scenarios is a set of n-tuples. In a given n-tuple, each element is a random variable that picks a realization of the fraction dose distribution. The elements are i.i.d. and referred to as T_f. To each element u in the set $\mathcal{U}$ of standard deviation scenarios corresponds a standard deviation σ_u.

The expectation under the probability distribution x is denoted by $\mathbb{E}_x$. When no subscript is used, the operation is taken with respect to the *a priori* probability distribution p or $q(\sigma)$, depending on the context.

TABLE I. Notation for the different types of uncertainties.

Uncertainty	Scenario set	Random variable	Distribution
Systematic	$\mathcal{S}$	S	p or π
Random	$\mathcal{T}$	T	$q(\sigma)$
Standard deviation	$\mathcal{U}$	U	p_σ or ρ

II.B. Notation

The set of feasible treatment parameters (spot weights, bixel weights, or machine settings) is denoted by $\mathcal{X}$. The total dose vector d is a function of $x \in \mathcal{X}$ and, depending on the uncertainties accounted for, $s \in \mathcal{S}$ or $t \in \mathcal{T}$ or both. It specifies the dose to each voxel v in the set $\mathcal{V}$ enumerating the voxels, and is the sum of n fraction doses. For a given ROI, the dose-to-volume D_y denotes the minimum dose level with an isodose volume containing y% of the ROI.

The vector of ones with dimension given by the context is denoted by e. Vector inequalities are to be understood componentwise. The shorthand y_+ is used for $\max\{y, 0\}$, the positive part of y.

II.C. Optimization functions

The methods for systematic errors can be used with general, nonlinear optimization functions. For a rigorous handling of random errors in a finite number of fractions, not all functions are computationally tractable. Many functions become tractable when the number of fractions is infinite, but this often leads to heterogeneous dose distributions.[4,12,16] In Secs. III and IV, it is shown how infinitely many fractions can be assumed without dose heterogeneity as a result. Nevertheless, for rigor to be maintained, the functions used in the present paper are restricted to uniform dose functions.

Given a set $\mathcal{R}$ of goals and for each goal $r \in \mathcal{R}$ an associated importance weight $w_r > 0$ and a prescribed dose level $\hat{d}_r$, the composite optimization function f is formulated

$$f(d) = \sum_{r\in\mathcal{R}} w_r \sum_{v\in\mathcal{V}_r} \Delta_{v,r}(d_v - \hat{d}_r)^2,$$

where $\mathcal{V}_r$ is the set of voxels of the ROI associated with goal $r \in \mathcal{R}$ and $\Delta_{v,r}$ is the relative volume of voxel v in ROI r, such that $\sum_{v\in\mathcal{V}_r} \Delta_{v,r} = 1$ for each $r \in \mathcal{R}$.

II.D. Accounting for systematic errors

Here, methods accounting for systematic errors only are formulated. For clarity, three special cases of methods in the considered class are formulated separately, followed by the general formulation.

II.D.1. Expected value optimization

The expected value optimization method minimizes the expected value of the objective and is the least conservative method in the considered class. It is formulated as the stochastic programming problem

$$\underset{x\in\mathcal{X}}{\text{minimize}} \quad \mathbb{E}[f(d(x, S))].$$

II.D.2. Worst case optimization

The worst case optimization (or "minimax") method minimizes the penalty of the worst scenario, with no regard to the probabilities of the scenarios. It is the most conservative method in the considered class. It is formulated as the minimax problem

$$\underset{x\in\mathcal{X}}{\text{minimize}} \quad \max_{s\in\mathcal{S}} f(d(x, s)).$$

II.D.3. Conditional value at risk optimization

The conditional value at risk (CVaR) is a coherent measure of risk introduced by Artzner *et al.*[14] and shown to be suitable for optimization by Rockafellar and Uryasev.[15] The CVaR function measures the mean tail loss, meaning the expected value of the fraction $0 < \alpha \leq 1$ of the worst scenarios, conditioned on that one of those scenarios will occur. It thus generalizes the expected value and worst case optimizations and allows for continuous scaling between the methods. The optimization problem takes the form

$$\begin{aligned} \underset{\lambda, x}{\text{minimize}} \quad & \lambda + \tfrac{1}{\alpha}\mathbb{E}[(f(d(x,S)) - \lambda)_+] \\ \text{subject to} \quad & x \in \mathcal{X}. \end{aligned} \tag{1}$$

When $\alpha = 1$, the problem corresponds to expected value optimization, whereas when $\alpha \leq \min_{s\in\mathcal{S}} p_s$ (assuming $p_s > 0$ for all $s \in \mathcal{S}$), it corresponds to worst case optimization.

II.D.4. Minimax stochastic programming

The three methods described above are special cases of the minimax stochastic formulation.[6,7] For our purposes, it can be formulated as

$$\underset{x\in\mathcal{X}}{\text{minimize}} \quad \max_{\substack{a \leq \pi \leq b \\ e^T\pi = 1}} \mathbb{E}_\pi[f(d(x,S))] \tag{2}$$

for vectors a and b satisfying $0 \leq a \leq b \leq e$ and $e^T a \leq 1 \leq e^T b$. Other linear restrictions on the probability distribution π may also be imposed.

Given x, the max function in the objective is a linear program in which $f(d(x, s))$ are known constants for $s \in \mathcal{S}$. This program is bounded and feasible, so, by strong duality for linear programming, its optimal value equals that of its dual. Substituting the dual for the max problem yields the equivalent formulation

$$\begin{aligned} \underset{\lambda,\mu,\nu,x}{\text{minimize}} \quad & \lambda + b^T\mu - a^T\nu \\ \text{subject to} \quad & \lambda + \mu_s - \nu_s \geq f(d(x,s)), \quad s \in \mathcal{S}, \\ & \mu, \nu \geq 0 \\ & x \in \mathcal{X}, \end{aligned} \tag{3}$$

which is better suited for optimization since the reformulation removes the discontinuities in the derivative of the objective. The λ of the CVaR problem (1) corresponds to the λ in (3), as explained in Appendix A.

Some special cases of parameters a and b are worth noting:

- $a = p$ or $b = p$: expected value optimization
- $a = 0$, $b = e$: worst case optimization
- $a = 0$, $b = \frac{1}{\alpha}p$: CVaR optimization with parameter α

II.E. Accounting for random errors

Because the number of scenarios in $\mathcal{T}$ grows exponentially with the number of fractions, minimax stochastic formulations for random errors are intractable in general. Still, they are tractable in special cases: Unkelbach and Oelfke[4] note that under the assumption of i.i.d. random errors, the expected value of the quadratically penalizing (uniform dose) optimization function under random errors can be easily computed. For a voxel $v \in \mathcal{V}$ with reference dose $\hat{d}$, the expected value of the uniform dose function is given by

$$\begin{aligned} \mathbb{E}[(d_v(x,T) - \hat{d})^2] = (\mathbb{E}[d_v(x,T_f^n)] - \hat{d})^2 \\ + \frac{1}{n}\mathrm{Var}(d_v(x,T_f^n)). \end{aligned} \tag{4}$$

Since the expectation and variance are taken with respect to the random variable T_f instead of T, the computational effort of evaluating this expression is independent of the number of fractions.

For other functions, it is seldom possible to reduce the exponential dependency on the number of fractions, but Jensen's inequality yields that the expectation of a maximum (and analogously for minimum) dose quadratic penalty to a voxel $v \in \mathcal{V}$ with reference dose $\hat{d}$ satisfies the inequalities

$$\begin{aligned} (\mathbb{E}[d_v(x,T_f^n)] - \hat{d})_+^2 &\leq \mathbb{E}[(d_v(x,T) - \hat{d})_+^2] \\ &\leq (\mathbb{E}[d_v(x,T_f^n)] - \hat{d})_+^2 \\ &\quad + \frac{1}{n}\mathrm{Var}(d_v(x,T_f^n)). \end{aligned}$$

When the number of fractions is infinite, the expectation of uniform, minimum, and maximum dose functions equal the respective functions applied to the expected dose. This approximation is frequently used to handle random errors. Without countermeasure, it is not always advisable, since the resulting dose distributions tend to be highly heterogeneous.[4,12,16] If the probability distribution used during the optimization is realized during treatment and the number of fractions is large, the alternating hot and cold spots in the heterogeneous dose distribution are likely to cancel each other out, but if the realized probability distribution is a different one, or if the number of fractions is not large enough, the hot and cold spots usually remain and the plan quality suffers.

One way of reducing the dose heterogeneity is to incorporate uncertainty in the probability distribution into the optimization. Hårdemark *et al.*[16] reduced the dose heterogeneity by optimizing the sum of the expected value objective for the standard deviation equal to zero and equal to its assumed maximum value. Unkelbach and Oelfke[17] also optimized such an expected value, but assumed that the random standard deviation was normally distributed.

In this paper, the uncertain random error standard deviation is accounted for by a minimax stochastic formulation, which generalizes the expected value methods used by previous authors. This formulation takes its inspiration from the method of Chan *et al.*,[12] who enforced an expected target dose greater than the prescription for all probability distributions within a given set. Here, the set of probability distributions is comprised of normal distributions for a number of different standard deviations. Maximization is performed over the probabilities of the different standard deviations before the expectation with respect to standard deviations and random errors is taken. Bounds on the probabilities for the different standard deviation can be used to adjust the conservativeness

of the method. The problem is formulated

$$\underset{x\in\mathcal{X}}{\text{minimize}} \overbrace{\max_{\substack{a_\sigma\le\rho\le b_\sigma\\ e^T\rho=1}} \mathbb{E}_\rho\left[\underbrace{\mathbb{E}_{q(\sigma_U)}\left[f(d(x,T))\right]}_{\text{Expectation w.r.t. random errors}}\right]}^{\text{Maximization over random standard deviation probability distribution}}, \quad \underbrace{\qquad\qquad}_{\text{Expectation w.r.t. random standard deviation}} \tag{5}$$

where the vectors a_σ and b_σ satisfy $0 \le a_\sigma \le b_\sigma \le e$ and $e^T a_\sigma \le 1 \le e^T b_\sigma$. This optimization problem can be reformulated as the problem for systematic errors, yielding a formulation much like (3). When $a_\sigma = b_\sigma = p_\sigma$, the expected value of the penalties under the different standard deviations is minimized; when $a_\sigma = 0$ and $b_\sigma = e$, the penalty under the worst standard deviation is minimized; and when $a_\sigma = 0$ and $b_\sigma = (1/\alpha)p_\sigma$, the CVaR with parameter α with respect to the standard deviations—i.e., the conditional expectation of the penalty over the α worst standard deviations—is minimized. As before, the dependency on T can be reduced to a dependency on T_f when uniform dose functions are used.

II.F. Combining systematic and random errors

Here, the formulations (2) and (5) are combined to handle the different types of uncertainty simultaneously. For each systematic error scenario, the minimax stochastically worst standard deviation is used in the expectation of the objective with respect to the random errors. Given these values for the systematic error scenarios, the minimax stochastically worst value is minimized. The problem is formulated as

$$\underset{x\in\mathcal{X}}{\text{minimize}} \overbrace{\max_{\substack{a\le\pi\le b\\ e^T\pi=1}} \mathbb{E}_\pi\left[\overbrace{\max_{\substack{a_\sigma\le\rho\le b_\sigma\\ e^T\rho=1}} \mathbb{E}_\rho\left[\underbrace{\mathbb{E}_{q(\sigma_U)}\left[f(d(x,S,T))\right]}_{\text{Expectation w.r.t. random errors}}\right]}^{\text{Maximization over random standard deviation probability distribution}}\right]}^{\text{Maximization over systematic error probability distribution}}, \quad \underbrace{\qquad\qquad\qquad\qquad}_{\text{Expectation w.r.t. random standard deviation}} \quad \underbrace{\qquad\qquad\qquad\qquad}_{\text{Expectation w.r.t. systematic errors}} \tag{6}$$

The outer minimax stochastic problem can be rewritten as before, yielding a formulation similar to (3) but with terms

$$\max_{\substack{a_\sigma\le\rho\le b_\sigma\\ e^T\rho=1}} \mathbb{E}_\rho[\mathbb{E}_{q(\sigma_U)}[f(d(x,s,T))]]$$

substituted for $f(d(x,s))$ in the constraints. These maxima can similarly be dualized, which introduces the additional auxiliary variables $\bar{\lambda}(s)$, $\bar{\mu}(s)$, and $\bar{\nu}(s)$ for $s \in \mathcal{S}$, where $\bar{\lambda}(s)$ is scalar whereas $\bar{\mu}(s)$ and $\bar{\nu}(s)$ are vectors of length $|\mathcal{U}|$. The problem turns into the two-level minimax stochastic problem

$$\begin{aligned}
\underset{\substack{\lambda,\mu,\nu,x,\\ \bar{\lambda},\bar{\mu},\bar{\nu}}}{\text{minimize}} \quad & \lambda + b^T\mu - a^T\nu \\
\text{subject to} \quad & \lambda + \mu_s - \nu_s \ge \bar{\lambda}(s) + b_\sigma^T\bar{\mu}(s) - a_\sigma^T\bar{\nu}(s), && s\in\mathcal{S},\\
& \bar{\lambda}(s) + \bar{\mu}_u(s) - \bar{\nu}_u(s) \ge \mathbb{E}_{q(\sigma_u)}[f(d(x,s,T))], && u\in\mathcal{U}, s\in\mathcal{S},\\
& \bar{\mu}(s), \bar{\nu}(s) \ge 0, && s\in\mathcal{S},\\
& \mu, \nu \ge 0 \\
& x \in \mathcal{X}.
\end{aligned} \tag{7}$$

When all uncertainties are handled by expected value optimization, the formulation can be simplified into a problem without constraints (besides $x \in \mathcal{X}$). When all uncertainties are handled by worst case optimization, the formulation can be simplified into a one-level problem with one constraint per scenario through removal of all auxiliary variables but λ. When all uncertainties are handled by CVaR optimization, the auxiliary variables ν and $\bar{\nu}(s)$ for $s \in \mathcal{S}$ can be removed. The formulations with many nonlinear constraints may make the problem more computationally demanding, depending on the optimization algorithm used. However, the main computational cost for this type of problem is usually the dose and the gradient calculations, which are required in equal number for all of the methods.

II.G. Patient geometry

The robust methods of the considered class can be applied to general clinical cases, as has been done previously for some special cases of the methods.[5,18] For illustratory purposes, a two-dimensional phantom geometry is studied here. The geometry is adapted as one slice of a C-shaped case modeled after that in the AAPM Task Group 119 Report[19] (but of larger outer target radius) and is shown in Fig. 1. Three beams with gantry angles 315°, 0°, and 45° are used. The ROIs are the C-shaped target, the circular organ at risk (OAR), and the circular external ROI. The case is discretized into 2×2 mm^2 voxels.

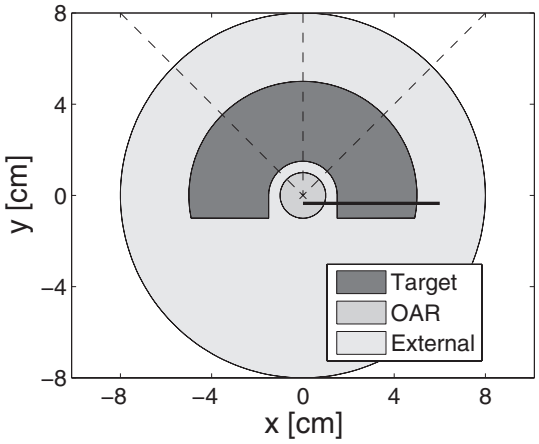


FIG. 1. The C-shaped geometry. The solid line indicates where the line doses are taken and the dashed lines indicate the beam directions. The radius of the inner arc of the target is 1.5 cm and that of the outer arc is 5 cm. The OAR has radius 1cm and the external ROI has radius 8 cm.

II.H. Computational study

A computational study in which IMPT was applied to the two-dimensional geometry was performed. Plans were optimized with the expected value, CVaR with parameter $\alpha = 0.5$, and worst case optimization methods. These methods handle the uncertainties with increasing conservativeness. First, the methods were applied for the case of only systematic errors; second, for the case of only random errors; and third, for the case of both systematic and random errors. As a reference, a conventional plan with a 5 mm margin as sole means of handling uncertainties was optimized. The results for the conventional plan are shown in Appendix B.

For all optimizations, the optimization functions used were the quadratic penalties on dose deviations described in Sec. II.C. The target had a normalized prescribed dose of 1, whereas the OAR and the external ROI had prescribed maximum doses of 0. The importance weights were 100, 10, and 1 for, respectively, the target, OAR, and external ROI. For systematic errors and for the combination of systematic and random errors, the effects of the importance weights were assessed by additional optimizations for ten different target weights, logarithmically spaced in [100, 1000].

The systematic errors were assumed to consist of range and setup errors. The range errors were modeled as uniform scalings of the patient densities with probability distribution taken from a discretized normal distribution with standard deviation 3.5%, corresponding to a shift of about 2.5 mm at the depth of the OAR. Absolute deviations larger than two standard deviations were cutoff and the distribution was renormalized. The setup uncertainties were assumed to be isotropic and with probability distribution in each axis direction taken from a discretized normal distribution with standard deviation 2.5 mm. A cutoff was applied such that the 95% most probable scenarios in the joint distribution were included and the distribution was renormalized.

Only setup errors were assumed to constitute the random errors, since range errors are likely systematic.[1] The random setup error probability distribution in each fraction and axis direction was taken from a discretized normal distribution but with the standard deviation a random variable uniformly distributed in [0, 5] mm. A cutoff was applied such that the 99% most probable scenarios in the joint distribution were included and the distribution was renormalized.

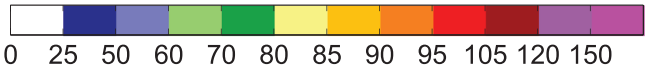


FIG. 2. Color table indicating the dose levels for the dose distributions. The tick labels denote percentage of the reference dose level, which is 1 for total doses and 0.5 for beam doses.

Five density scalings and five random error standard deviations were included. A grid of shifts with 2.5 mm spacing was used as basis for the setup scenarios, and the most probable scenarios within the cutoffs given above were included. This resulted in 89 scenarios for systematic errors and 109 scenarios for random errors. Since this study deals with the aims of the methods rather than the accuracy of the scenario discretization, the same scenarios were used in the optimization and in the evaluation.

Proton pencil beam kernels were calculated with the pencil beam dose algorithm of the RayStation treatment planning system version 2.6 (RaySearch Laboratories, Stockholm, Sweden). The line spacing was 2.5 mm and the energy spacing was 2.5 mm in water. The optimizations were performed in MATLAB version 7.9 using the sequential quadratic programming solver SNOPT version 7.2.[20]

The methods are assessed by comparison of their total and beam dose distributions, line doses, dose-volume histogram (DVH) families over simulated realizations of errors, trade-off curves, and dose-probability histograms (DPHs). The trade-off curves show how the target D_{95} is traded for the OAR D_{25} in the worst (lowest target D_{95}, highest OAR D_{25}), mean, and best (highest target D_{95}, lowest OAR D_{25}) case over a number of simulated treatments when the importance weight of the target optimization function is changed. The DPHs show the probability of achieving a given dose-to-volume for a number of dose-to-volume levels (D_{95} and D_{95} for the target, D_5 and D_2 for the OAR), where each level is represented by a curve. A DPH provides information such as "with 90% probability, D_{95} of the target will be 0.97 or higher." The DPHs and trade-off curves are based on 1000 simulated treatments subject to the considered uncertainties, except when only systematic errors are handled, in which case the 89 systematic error scenarios are used. The color table indicating the levels of the dose distributions is displayed in Fig. 2.

III. RESULTS

III.A. Systematic errors

When considering only systematic range and setup errors, the formulation (3) was applied. Total doses resulting

from the robust methods are displayed in Fig. 3. As the conservativeness of the method increases, the high-dose region outside the target is extended and the dose in a semicircle around the OAR is escalated. The total doses in the perturbed scenario display less underdosage for the worst case method than for the other methods.

The beam doses in Fig. 4 show that the robust methods account for interplay between beams by using small beam dose gradients. The oblique beams patch smoothly, which makes for low sensitivity to relative mispositioning. The worst case method uses the 0° beam less than the other methods and the CVaR method uses that beam more.

The DVH families in Fig. 5 show that the target coverage robustness increases with the conservativeness of the method. For the worst case method, the best case OAR dose is slightly worse than in the other methods.

Figure 6(a) displays the trade-off curves between target coverage and OAR sparing, as resulting when the importance weight of the target optimization function is increased. It shows that the worst case trade-off improves with the conservativeness of the method. The mean and best case trade-off curves are more similar than the worst case curves, but the worst case optimization leads to worse mean and best case trade-off than the other methods. The difference in best case trade-off is especially large for low target weights. It is notable that the CVaR method leads to a better mean trade-off than the other methods. The DPHs in Fig. 6(b) also show that the target coverage increases with the conservativeness, and that the worst case method often leads to higher OAR doses than the other methods, although it has lower probabilities for the highest OAR doses.

III.B. Random errors with fixed probability distribution

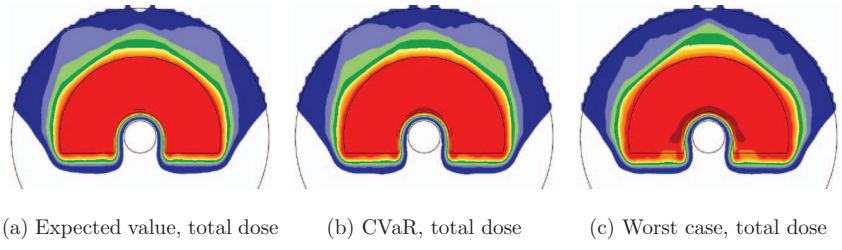
When only random errors with an assumedly known probability distribution were handled, an objective function consisting of terms like (4) was used. The random error standard deviation was assumed to be 5 mm. Figure 7 shows the resulting total dose distributions for the number of fractions n in $\{30, \infty\}$ and the line doses for n in $\{1, 5, 30, \infty\}$. The dose distributions become increasingly undulated as the number of fractions increases.

III.C. Random errors with uncertain standard deviation

To reduce dose variability, uncertainty in the standard deviation of the random errors was included in the optimization. The methods were formulated according to (5). The standard deviation was assumed to be uniformly distributed in [0, 5] mm. It was assumed that there were no systematic errors.

For all methods, the resulting dose distributions for $n = 30$ and $n = \infty$ were very similar (root mean square differences below 0.02). The case with $n = \infty$ has no variance terms that penalize heterogeneity and is thus more likely to result in a heterogeneous plan. This case is therefore the main focus of the subsequent presentation. Total and beam doses of the CVaR method for $n = \infty$ are displayed in Fig. 8. The doses of the other robust methods were similar, with root mean square differences from the CVaR doses below 0.02 for total as well as beam doses. The corresponding values for the conventional plan, shown in Appendix B, were an order of magnitude larger. Accounting for uncertainty in the standard deviation resulted in much reduced dose variance compared to the case with fixed standard deviation.

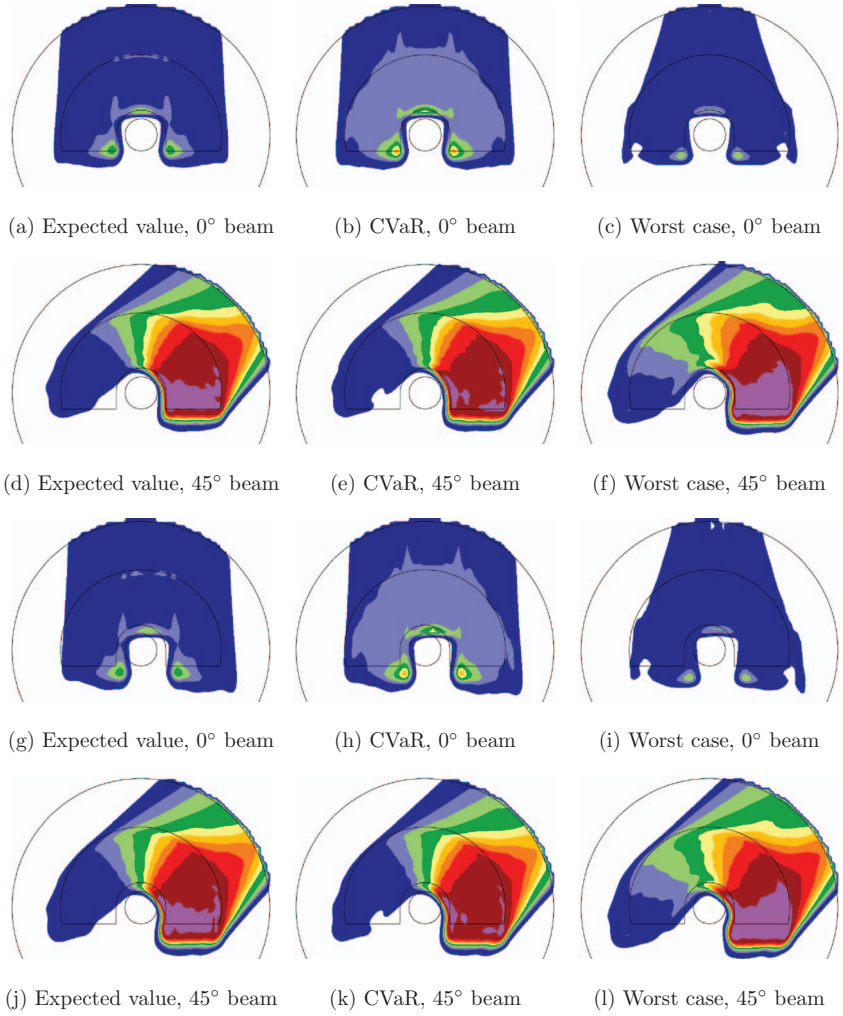
DVH families for 100 realizations of random standard deviations and random errors in 30 fractions (but with $n = \infty$ during optimization) for the robust methods are shown in Fig. 9. Like the dose distributions, the DVH families are quite similar between the methods, but the target coverage robustness increases with the conservativeness. This difference is more readily seen in the DPHs, shown in Fig. 10. The DPHs,



(a) Expected value, total dose (b) CVaR, total dose (c) Worst case, total dose

(d) Expected value, total dose (e) CVaR, total dose (f) Worst case, total dose

FIG. 3. Total doses for the robust methods for systematic errors. (a)–(c) Nominal scenario; and (d)–(f) isocenters shifted 0.5 cm to the right and density two standard deviations lower than measured.



(a) Expected value, 0° beam (b) CVaR, 0° beam (c) Worst case, 0° beam

(d) Expected value, 45° beam (e) CVaR, 45° beam (f) Worst case, 45° beam

(g) Expected value, 0° beam (h) CVaR, 0° beam (i) Worst case, 0° beam

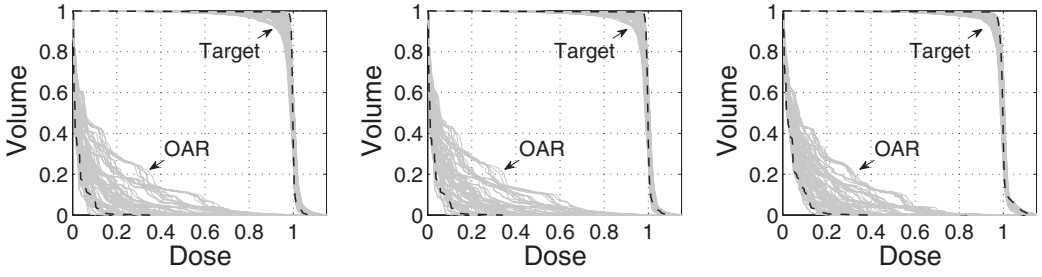
(j) Expected value, 45° beam (k) CVaR, 45° beam (l) Worst case, 45° beam

FIG. 4. Beam doses for the robust methods for systematic errors. (a)–(f) Nominal scenario; and (g)–(l) isocenters shifted 0.5 cm to the right and density two standard deviations lower than measured.

moreover, show that the OAR doses decrease as the conservativeness increases.

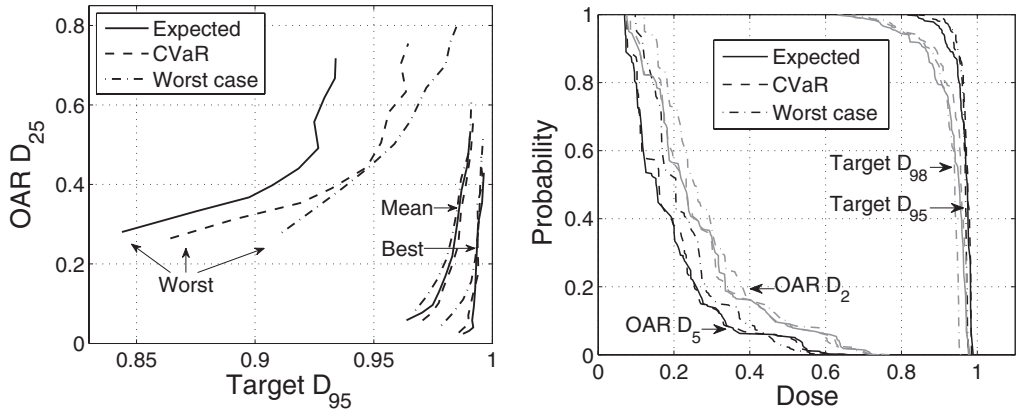
Line doses are displayed in Fig. 11. All methods keep a horn at the left-hand side for $n > 1$: while reducing the dose variance, some dose escalation in a semicircle around the OAR is kept. There is also a slight dose escalation along the proximal target periphery. The height of the horn increases slightly with the conservativeness of the method.

For the worst case method, only the constraints for the extreme scenarios with standard deviations 0 and 5 mm were satisfied with equality in the optimum. The impact of the scenario with zero standard deviation increased with the number



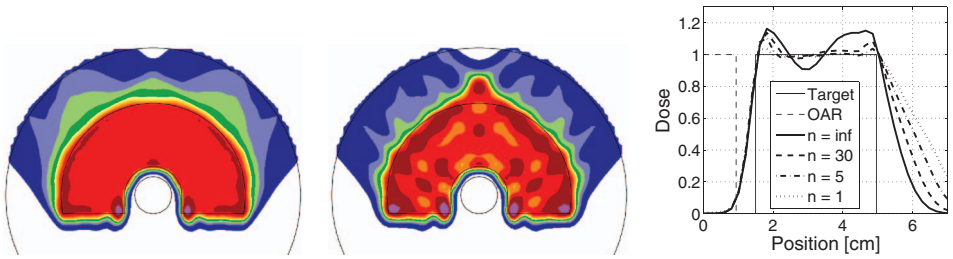
(a) Expected value (b) CVaR (c) Worst case

FIG. 5. DVH families for the robust methods for systematic errors over the 89 systematic error scenarios. The dashed lines correspond to the nominal scenario DVHs.



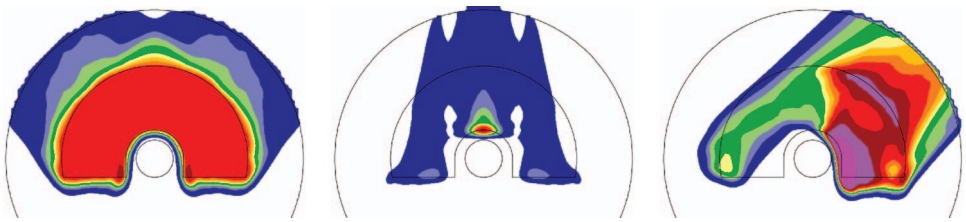
(a) (b)

FIG. 6. (a) Worst (lowest target D_{95}, highest OAR D_{25}), mean, and best (highest target D_{95}, lowest OAR D_{25}) case trade-off curves for the target importance weight in [100, 1000]; and (b) DPHs. Both figures are resulting from the robust methods for systematic errors evaluated over the 89 systematic error scenarios.



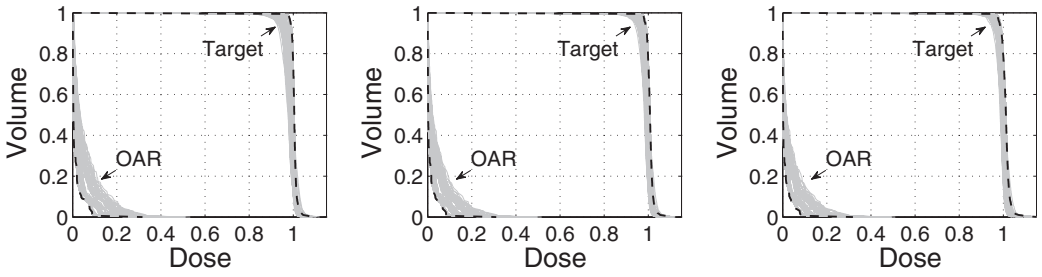
(a) Expected value, $n = 30$ (b) Expected value, $n = \infty$ (c) Expected value, line doses

FIG. 7. Total nominal scenario doses for the expected value optimization for random errors with fixed standard deviation for the number of fractions (a) $n = 30$ and (b) $n = \infty$; and (c) line doses for n in $\{1, 5, 30, \infty\}$ taken along the line shown in Fig. 1.



(a) CVaR, total dose (b) CVaR, 0° beam (c) CVaR, 45° beam

FIG. 8. Nominal scenario total and beam doses for the CVaR optimization for random errors with uncertain standard deviation. The other robust methods resulted in similar dose distributions (root mean square differences from CVaR below 0.02).



(a) Expected value (b) CVaR (c) Worst case

FIG. 9. DVH families for the robust methods for random errors with uncertain standard deviation over 100 realizations of random standard deviations and random errors in 30 fractions. The optimizations were performed with $n = \infty$. The dashed lines correspond to the nominal scenario DVHs.

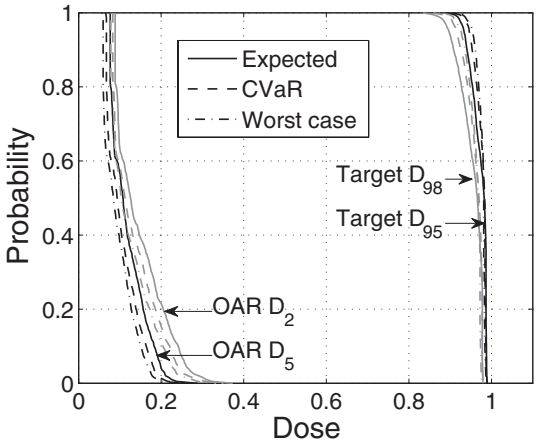
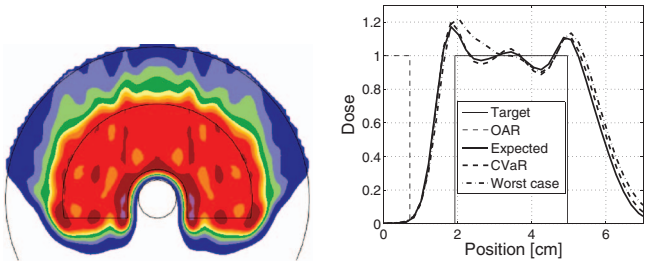


FIG. 10. DPHs for the robust methods for random errors with uncertain standard deviation based on 1000 simulations of random standard deviations and random errors in 30 fractions. The optimizations were performed with $n = \infty$.



(a) CVaR, total dose (b) Line doses

FIG. 12. (a) Total nominal scenario dose for the CVaR optimization for systematic errors and random errors with fixed standard deviation. The number of fractions $n = \infty$. The other robust methods resulted in similar heterogeneous dose distributions, which is reflected in inset (b), the line doses taken along the line shown in Fig. 1.

of fractions, which was determined from the marginal costs (Lagrange multipliers) of the constraints.

III.D. Systematic errors and random errors with fixed probability distribution

To account for systematic errors as well as random errors with an assumedly known probability distribution, the formulation (3) was applied to functions consisting of terms like (4). The random error standard deviation was assumed to be 5 mm. Figure 12 shows the resulting total dose distribution and line doses. As in the case with only random errors, the dose distributions are highly undulated.

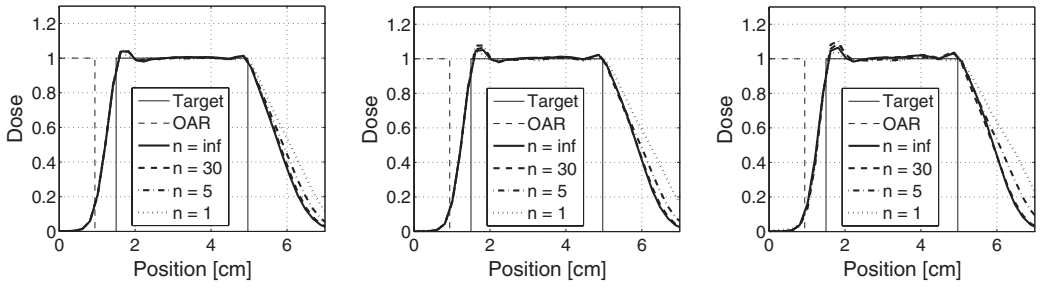
III.E. Systematic errors and random errors with uncertain standard deviation

When accounting for systematic errors and random errors with uncertain standard deviation, the formulation (7) was used. The robust methods were applied to their full extents. Thus, when the expected value was optimized, the systematic errors as well as the uncertain standard deviation were handled by expectation, and analogously for the CVaR and worst case methods. The random errors were always handled by expectation under the approximation of infinitely many fractions.

The resulting total and beam dose distributions are shown in Figs. 13 and 14. Accounting for the random standard deviation reduces the undulations that were present for the case with fixed standard deviation. The doses share most characteristics with the case of only systematic errors: the high-dose regions extend outside the target and there are dose escalations in a semicircle around the OAR. There is, however, also a small increase in dose along the proximal target periphery, as when only random errors with uncertain probability distribution were accounted for. All these effects increase with the conservativeness of the method.

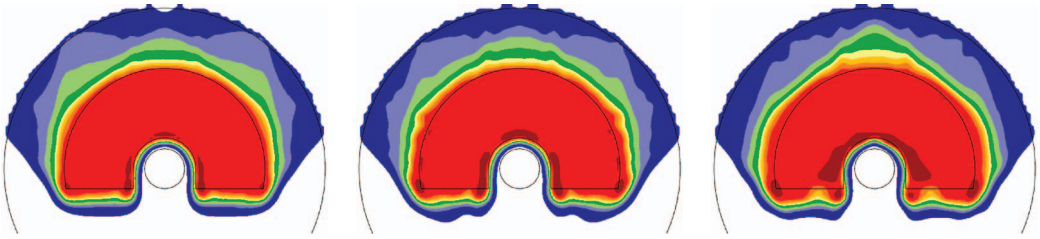
DVH families for 100 realizations of systematic errors, random standard deviations, and random errors in 30 fractions (but with $n = \infty$ during optimization) are displayed in Fig. 15. The worst case method results in more robust target coverage and, for most volumes, higher OAR doses than the other methods.

Trade-off curves between target coverage and OAR sparing for the different methods are shown in Fig. 16(a). The characteristics are the same as when only systematic errors were included: the worst case trade-off curves improve with the conservativeness of the method, worst case optimization



(a) Expected value (b) CVaR (c) Worst case

FIG. 11. Nominal scenario line doses taken along the line shown in Fig. 1 for the robust methods for random errors with uncertain standard deviation for the number of fractions n in $\{1, 5, 30, \infty\}$.



(a) Expected value, total dose (b) CVaR, total dose (c) Worst case, total dose

FIG. 13. Total nominal scenario doses for the robust methods for systematic errors and random errors with uncertain standard deviation. The number of fractions $n = \infty$.

leads to worse mean and best case trade-off than the other methods, and CVaR leads to the best mean trade-off.

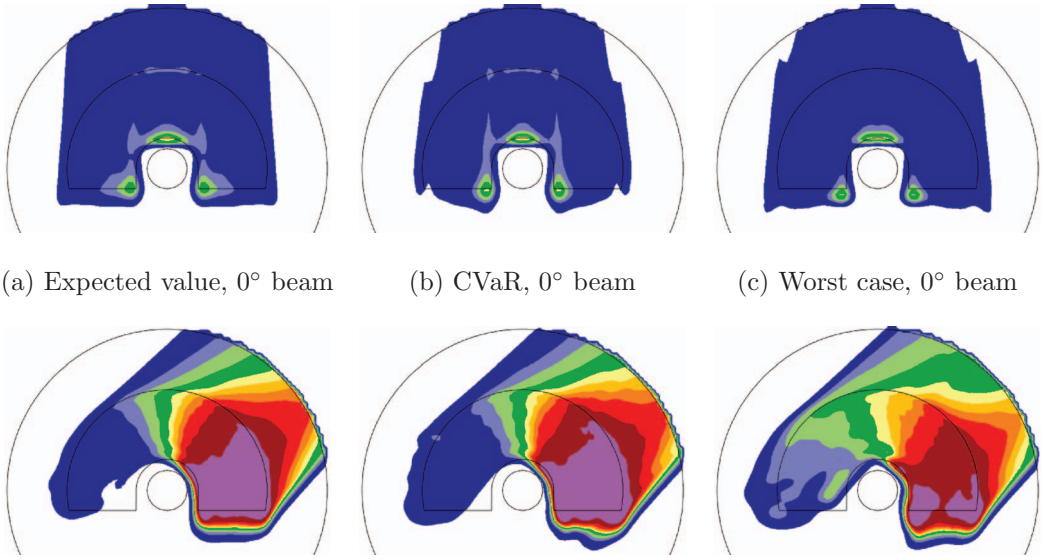
Like the dose distributions, the DVH families are reminiscent of those resulting when only systematic errors were accounted for. Therefore, DPHs with respect to systematic errors, random standard deviations, and random errors were computed not only for the methods accounting for these errors, but also for the methods accounting for systematic errors only. The DPHs are shown in Figs. 16(b) and 16(c). The target coverage robustness is better for the methods accounting for the different types of uncertainty than for the methods accounting for systematic errors only, as is the OAR sparing. Aside from that, similar observations can be made in both insets: the target coverage robustness increases with the conservativeness, and that the worst case method often leads to higher OAR doses than the other methods, although it has lower probabilities for the highest OAR doses.

Line doses for the robust methods are presented in Fig. 17. They are more homogeneous than when the standard deviation was assumed fixed, but small horns remain. The worst case method results in larger horns than the other methods and moreover, due to its prevalent use of the oblique beams, slower fall-off at the right-hand side.

IV. DISCUSSION

For complex patient cases, the realization of errors often drastically changes the dose distribution, which makes margins of any size and shape inadequate to achieve robustness. The incorporation of information about the uncertainties into the optimization together with the utilization of robust optimization techniques enable the optimization algorithm to account for these changes.

For all types of uncertainties, the target coverage robustness increased with the conservativeness of the applied method. When systematic errors were included, worst case optimization generally resulted in worse OAR doses than the other methods. CVaR resulted in better target coverage than expected value minimization, and at the same time similar OAR sparing. For random errors, both target coverage and OAR sparing improved with the conservativeness of the method. This shows that it can be worthwhile to use more



(a) Expected value, 0° beam (b) CVaR, 0° beam (c) Worst case, 0° beam

(d) Expected value, 45° beam (e) CVaR, 45° beam (f) Worst case, 45° beam

FIG. 14. Nominal scenario beam doses when random errors with uncertain standard deviation are handled by the robust methods. The number of fractions $n = \infty$.

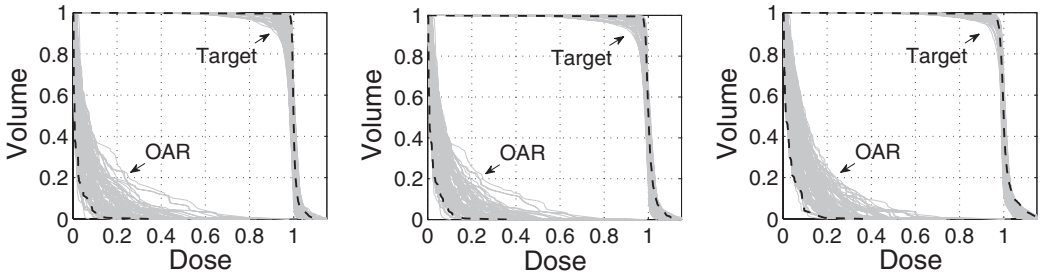


FIG. 15. DVH families for the robust methods for systematic errors and random errors with uncertain standard deviation over 100 realizations of systematic errors, random standard deviations, and random errors in 30 fractions. The optimizations were performed with $n = \infty$. The dashed lines correspond to the nominal scenario DVHs.

conservative methods than the expected value method to account for errors in IMPT.

The robust methods for systematic errors all lead to smooth beam doses. This makes the total dose robust with respect to the beam doses shifting relative to each other, which is an effect of errors. For random errors, the multiple fractions make smooth beam fraction doses less important than for systematic errors: when the random errors are small, large beam dose gradients pose little problem and when the random errors are large, the beam doses are smoothened over the multiple fractions.

The trade-off curves for systematic errors showed that the expected value method resulted in unnecessarily high worst case OAR doses for a given worst case target dose. The underlying reason is that when the errors are handled by expectation, the penalties of the different scenarios are multiplied by their respective probabilities. For some scenarios, the probability-adjusted target penalty will be balanced by the penalties of dose-reducing goals. For volumes outside the target, the aim of the expected value optimization is thus to deliver doses between the prescribed target dose level and zero dose. These in-between doses contribute only marginally to the target coverage when systematic errors are realized. It is more appropriate to aim at either zero dose (for better OAR sparing) or a dose closer to the target prescription (for more robust target coverage). The more conservative methods do so to a higher degree.

The common technique of handling random errors by applying the optimization functions to the expectation of the dose was observed leading to alternating hot and cold spots. Incorporating uncertainty in the probability distribution of the random errors into the optimization resulted in more homogeneous dose distributions. Even if the standard deviations of the random errors are known, it may still be beneficial to incorporate other standard deviations in the optimization to achieve homogeneous fraction doses.

When uncertainties in the probability distributions of the random errors were accounted for (more specifically, when the nominal scenario was included), the difference between the number of fractions $n = 30$ and $n = \infty$ was reduced. This indicates that the assumption of infinitely many fractions may be viable provided uncertainty in the probability distribution of the random errors is handled. Under the approximation of infinitely many fractions, the common minimum and maximum dose and DVH functions can be used in a robust setting, since they are then to be applied to the expected dose. Because expectation is a linear operation, the expected dose is generally less costly to compute than the dose variance.

That only the constraints of the scenarios with extreme standard deviations were tight in the optimum of the worst case method for uncertain standard deviations suggests a computationally inexpensive approximation of the worst case method. For all cases studied here, the same solution would be achieved if only the extreme scenarios were included.

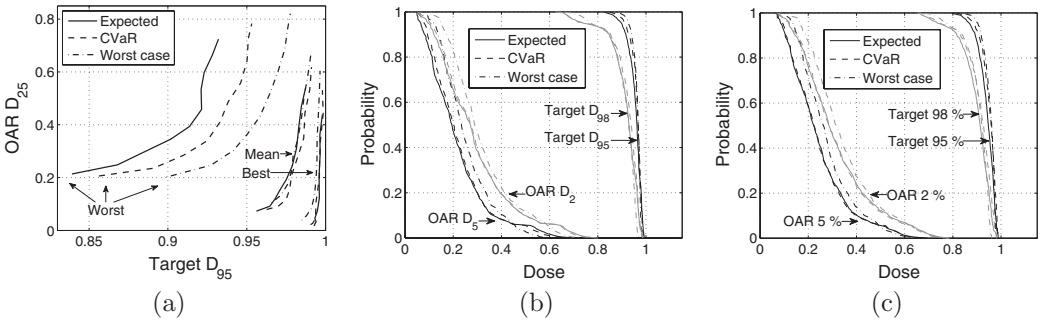


FIG. 16. (a) Worst (lowest target D_{95}, highest OAR D_{25}), mean, and best (highest target D_{95}, lowest OAR D_{25}) case trade-off curves of the methods accounting for systematic errors and random errors with uncertain standard deviation and $n = \infty$, for the target importance weight in [100, 1000]; (b) DPHs of the methods accounting for systematic errors and random errors with uncertain standard deviation and $n = \infty$; and (c) DPHs of the methods accounting for systematic errors only. The three figures are resulting from the robust methods evaluated over 1000 simulated treatments of systematic errors, random standard deviations, and random errors in 30 fractions.

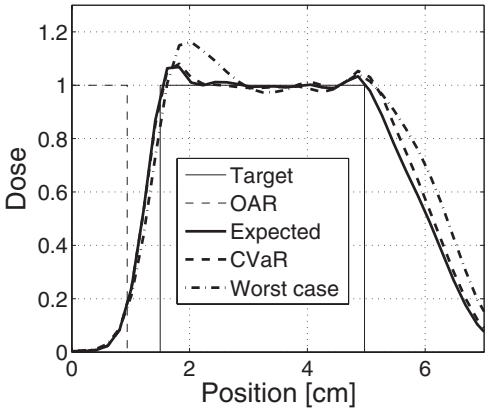


FIG. 17. Nominal scenario line doses taken along the line shown in Fig. 1 for the robust methods for systematic errors and random errors with uncertain standard deviation. The number of fractions $n = \infty$.

Chan *et al.*[13] showed this to be necessarily true for a one-dimensional geometry.

A problem with worst case optimization is that it neglects all scenarios but the ones with highest objective value. In optimum, it may thus be possible to reduce the objective values of other, inactive, scenarios without increasing the objective values of the worst scenarios. Similarly, CVaR optimization neglects the $1 - \alpha$ best scenarios. It may therefore be beneficial to add low-weighted expected value terms to the objectives of these methods, which would give the optimization incentive to always reduce the objective values of all, and not just the worst, scenarios.

It is worth noticing that another reason lies behind the dose escalations around the OAR resulting when only systematic errors were accounted for, shown in Fig. 3, than that when only random errors were accounted for, shown in Fig. 8 and also reported by other authors.[3,4,12] For systematic errors, the dose is escalated to build sharper penumbrae against the OAR (utilizing that the fall-off rate of Gaussian functions increases with the distance from the peak), whereas for random errors, the dose is escalated to ensure that underdosage in one fraction is likely to be compensated in another.

V. CONCLUSION

Accounting for uncertainties is essential to achieve robust radiation therapy treatment plans for complex cases. A class of robust optimization methods doing so by utilizing more information than conventional in the optimization was studied. The class includes expected value, CVaR, and worst case optimization. These methods were used to optimize IMPT plans for a two-dimensional C-shaped phantom subject to systematic errors and random errors with uncertain standard deviation. The target coverage robustness increased with the conservativeness of the method. For systematic errors, the OAR doses for most volumes increased with the conservativeness, and for random errors with uncertain distribution, they decreased. The inadequacy of handling random errors by optimizing the expected dose was highlighted, and it was shown that accounting for uncertainties in the probability distribution of the random errors provides a possible remedy.

ACKNOWLEDGMENTS

The author thanks Anders Forsgren and Björn Hårdemark for valuable comments on the paper. The research was supported by the Swedish Research Council (VR).

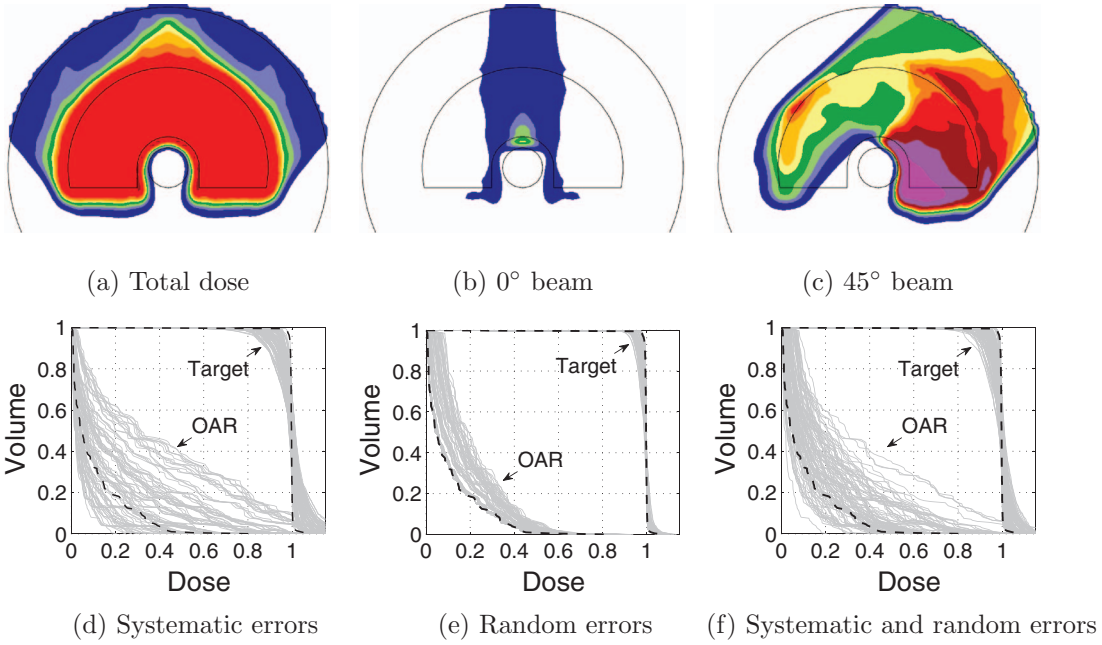


FIG. 18. Doses and DVH families for the conventional method with a 5 mm margin. (a) Total dose; (b) and (c) beam doses. (d) DVH family over the 89 systematic error scenarios; (e) over 100 realizations of random standard deviations and random errors in 30 fractions; and (f) over 100 realizations of systematic error as well as random standard deviations and random errors in 30 fractions. The dashed lines correspond to the nominal scenario DVHs.

APPENDIX A: CVAR AS A MINIMAX STOCHASTIC PROGRAM

The variables λ of the CVaR formulation (1) and the minimax stochastic formulation (3) have the same meaning: When the expectation of the positive part operator in (1) is handled by the introduction of additional variables μ, an equivalent CVaR formulation is yielded according to the following:

$$\begin{aligned} \underset{\lambda,\mu,x}{\text{minimize}} \quad & \lambda + \tfrac{1}{\alpha} p^T \mu \\ \text{subject to} \quad & \mu_s \geq f(d(x,s)) - \lambda, \quad s \in \mathcal{S}, \\ & \mu \geq 0 \\ & x \in \mathcal{X}. \end{aligned}$$

Here, λ and μ can be identified with λ and μ in (3) with $a = 0$ and $b = (1/\alpha)p$.

APPENDIX B: CONVENTIONAL PLANNING

The robust methods may be contrasted to conventional planning with a 5 mm margin. Its total and beam doses as well as DVH families are shown in Fig. 18.

[a]Electronic addresses: albfre@kth.se and albin.fredriksson@raysearchlabs.com

[1]A. J. Lomax, "Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: The potential effects of calculational uncertainties," Phys. Med. Biol. **53**(4), 1027–1042 (2008).
[2]A. J. Lomax, "Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: The potential effects of inter-fraction and inter-field motions," Phys. Med. Biol. **53**(4), 1043–1056 (2008).
[3]J. Löf, B. K. Lind, and A. Brahme, "Optimal radiation beam profiles considering the stochastic process of patient positioning in fractionated radiation therapy," Inverse Probl. **11**(6), 1189–1209 (1995).
[4]J. Unkelbach and U. Oelfke, "Inclusion of organ movements in IMRT treatment planning via inverse planning based on probability distributions," Phys. Med. Biol. **49**(17), 4005–4029 (2004).
[5]A. Fredriksson, A. Forsgren, and B. Hårdemark, "Minimax optimization for handling range and setup uncertainties in proton therapy," Med. Phys. **38**(3), 1672–1684 (2011).
[6]J. Dupačová, "Minimax stochastic programs with nonseparable penalties," In *Optimization Techniques, Lecture Notes in Control and Information Sciences* (Springer-Verlag, Berlin, 1980), pp. 157–163.
[7]A. Shapiro and A. Kleywegt, "Minimax analysis of stochastic problems," Optim. Methods Software **17**, 523–542 (2002).
[8]W. Chen, J. Unkelbach, A. Trofimov, T. Madden, H. Kooy, T. Bortfeld, and D. Craft, "Including robustness in multi-criteria optimization for intensity-modulated proton therapy," Phys. Med. Biol. **57**(3), 591–608 (2012).
[9]A. Ben-Tal and A. Nemirovski, "Robust convex optimization," Math. Oper. Res. **23**(4), 769–805 (1998).
[10]D. Bertsimas and M. Sim, "The Price of Robustness," Oper. Res. **52**(1), 35–53 (2004).
[11]M. Chu, Y. Zinchenko, S. G. Henderson, and M. B. Sharpe, "Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty," Phys. Med. Biol. **50**(23), 5463–5477 (2005).
[12]T. C. Y. Chan, T. Bortfeld, and J. N. Tsitsiklis, "A robust approach to IMRT optimization," Phys. Med. Biol. **51**(10), 2567–2583 (2006).
[13]T. C. Y. Chan, J. N. Tsitsiklis, and T. Bortfeld, "Optimal margin and edge-enhanced intensity maps in the presence of motion and uncertainty," Phys. Med. Biol. **55**(2), 515–533 (2010).
[14]P. Artzner, F. Delbaen, J. M. Eber, and D. Heath, "Coherent Measures of Risk," Math. Finance **9**(3), 203–228 (1999).
[15]R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," J. Risk **2**, 21–41 (2000).
[16]B. Hårdemark, J. Andersson, A. Liander, H. Rehbinder, and J. Löf, "Robust inverse treatment planning considering organ motion," in *Proceedings of the American Association of Physicists in Medicine (AAPM) 46th Annual Meeting (poster)*, 1715, 2004.
[17]J. Unkelbach and U. Oelfke, "Incorporating organ movements in inverse planning: Assessing dose uncertainties by bayesian inference," Phys. Med. Biol. **50**(1), 121–139 (2005).
[18]J. Unkelbach, T. Bortfeld, B. C. Martin, and M. Soukup, "Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning," Med. Phys. **36**(1), 149–163 (2009).
[19]G. A. Ezzell, J. W. Burmeister, N. Dogan, T. J. LoSasso, J. G. Mechalakos, D. Mihailidis, A. Molineu, J. R. Palta, C. R. Ramsey, B. J. Salter, J. Shi, P. Xia, N. J. Yue, and Y. Xiao, "IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119," Med. Phys. **36**(11), 5359–73 (2009).
[20]P. E. Gill, W. Murray, and M. A. Saunders, "SNOPT: An SQP algorithm for large-scale constrained optimization," SIAM Rev. **47**(1), 99–131 (2005).