Submitted to *INFORMS Journal on Computing*
manuscript (Please, provide the mansucript number!)

Robust optimization of dose-volume metrics for prostate HDR-brachytherapy incorporating target- and OAR volume delineation uncertainties

Marleen Balvert, Dick den Hertog
Department of Econometrics and Operations Research/Center for Economic Research (CentER), Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands, m.balvert@tilburguniversity.edu, d.denhertog@tilburguniversity.edu

Aswin L. Hoffmann
Institute of Radiooncology, Helmholtz-Zentrum Dresden-Rossendorf, Dresden, Germany
Department of Radiotherapy and Radiooncology, University Hospital Carl Gustav Carus at the Technische Universität Dresden, Dresden, Germany
Department of Radiation Oncology (MAASTRO), GROW School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, The Netherlands, aswin.hoffmann@oncoray.de

In radiation therapy planning, uncertainties in target volume definition yield a risk of underdosing the tumor. The classical way to prevent this in the context of external beam radiotherapy (EBRT) has been to expand the clinical target volume (CTV) with an isotropic margin to obtain the planning target volume (PTV). However, the EBRT-based PTV concept is not directly applicable to brachytherapy (BT) since it can lead to undesirable dose escalation (Tanderup et al. 2010). This work presents a treatment plan optimization model that uses worst-case robust optimization to account for delineation uncertainties in interstitial high-dose-rate BT of the prostate. A scenario-based method was developed that handles uncertainties in index sets. Heuristics were included to reduce the calculation times to acceptable proportions. The approach was extended to account for delineation uncertainties of an OAR as well. The method was applied on data from prostate cancer patients, and evaluated in terms of commonly used dosimetric performance criteria for the CTV and relevant organs at risk. The robust optimization approach was compared against the classical PTV margin concept and against a scenario-based CTV margin approach. The results show that both the scenario-based margin and the robust optimization method are well capable of reducing the risk of underdosage to the tumor. As expected, the scenario-based CTV margin approach leads to dose escalation within the target, whereas this could be prevented with the robust model. For cases where rectum sparing was a binding restriction, including uncertainties in rectum delineation in the planning model led to a reduced risk of a rectum overdose, and in some cases to reduced target coverage.

Key words: brachytherapy, delineation uncertainties, mixed integer linear optimization, robust optimization, treatment planning

1. Introduction

Cancer can be treated by surgery, chemotherapy, radiotherapy, or a combination of these modalities. For deep-seated solid tumors, radiotherapy is an adequate treatment option as ionising radiation can penetrate through healthy tissues to reach the tumor. Radiation therapy can either be delivered by external beam radiotherapy (EBRT) or by brachyther-

apy (BT). With EBRT, radiation coming from an external source is pointed at the tumor, while with BT a small radio-active source is placed inside or close to the tumor.

For prostate cancer, which is the most common type of cancer among men in the Western world, interstitial high-dose rate brachytherapy (HDR-BT) with a temporary implant has been shown to be an adequate treatment (Yamada et al. 2012). Typically, a template containing a large number of evenly spaced holes is placed in front of the patient's perineum while he is under anesthesia in dorsal position. Depending on the dimensions of the prostate, around 15 up to 20 of these holes are selected for implanting a hollow catheter into the prostate. After implantation of all needles, a remote afterloader device advances a ^{192}Ir source through the needles in a successive way. In each catheter, the source stops at predetermined locations (dwell positions) inside the target volume for a predetermined amount of time (dwell time) in order to deposit a sufficiently high dose to the tumor. Directly after irradiation, the source is removed from the patient and stored in the afterloader device for future use.

Radiotherapy inevitably results in exposure of healthy tissues surrounding the tumor. The spatial distribution of catheters and dwell positions together with the dwell time distribution determine the shape and magnitude of the dose distribution. The goal of treatment planning is to determine the number and locations of catheters together with the dwell times such that the tumor receives a sufficiently high dose to sterilize the tumorous cells while limiting the dose exposure to surrounding organs at risk (OARs) as much as possible to minimize the risk of side-effects. The problem of designing a treatment plan for HDR-BT lends itself to be formulated as a mathematical optimization problem (De Boeck et al. 2014).

Prior to treatment planning a scan of the patient's anatomy is made on which the target volume and the OARs are delineated as structure sets (Figure 1). These delineations are subject to intra- and inter-observer variability, i.e., the same observer does not draw identical contours for the same individual case, and different observers produce different delineations for an identical case, respectively (see e.g. Villeirs et al. (2005) and De Brabandere et al. (2012)). This implies uncertainty in the location and shape of the delineated structures, and thus uncertainty in the volumes to be irradiated and the volumes to be spared. In order to numerically optimize and evaluate a dose distribution, these structures are discretized into finite sets of small volume elements, that are considered as dose calculation points. The dose deposited in each calculation point is the superposition of the dose rate contributions from all the dwell positions weighted by their respective dwell times. Uncertainties in the delineations hence translate into uncertainty in whether or not a calculation point belongs to a certain structure (Figure 2). This implies that there is uncertainty in the index sets of the optimization model. So far, optimization methods have not dealt with this type of uncertainty. Therefore, the aim of the current work is to develop an optimization method that is robust against uncertainty in index sets, and consequently can be applied for robust optimization of HDR-BT dose distributions incorporating delineation uncertainties.

1.1. Review of methods accounting for uncertainties in treatment planning from clinical practice and literature

The classical way to deal with geometrical uncertainties in EBRT planning is to apply a margin around the tumor volume such that a sufficiently large volume receives the therapeutic dose that was prescribed (e.g. Van Herk et al. 2002, Van Herk 2004). A more recent

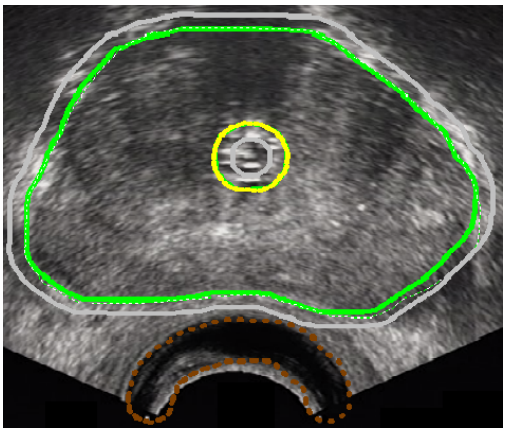

Figure 1 **A transrectal ultrasound image with delineated target volume (green) extended with an isotropic margin (gray), rectum (brown) and urethra (yellow). The gray delineation is a safety margin accounting for uncertainties, see Section 1.1.1.**

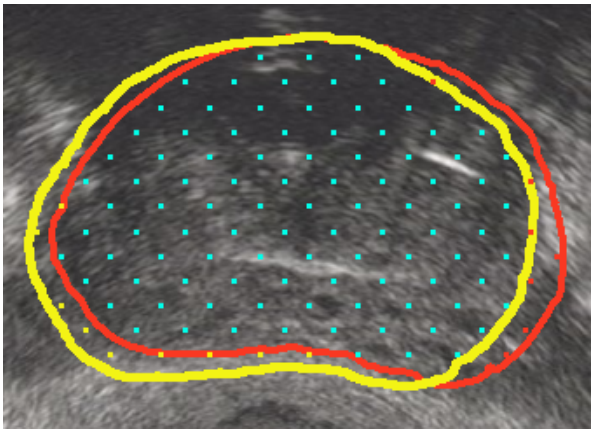

Figure 2 **Two delineations of a prostatic target volume based on a transversal ultrasound imaging scan. Both delineations yield a different set of calculation points residing in the structure: the blue points reside in both structures, the yellow and red points only in the yellow and red delineation, respectively.**

approach is to use computational methods to numerically account for uncertainties during the treatment planning optimization process. Methods like stochastic programming (e.g. Unkelbach and Ulfke 2004, Bohoslavsky et al. 2013) and worst-case robust optimization (e.g. Chan et al. 2006, Bortfeld et al. 2008, Fredriksson 2013) have been suggested for this task.

1.1.1. Margin approach According to international consensus guidelines published in the ICRU 62 report (International Commission on Radiation Units and Measurements 1999), uncertainties in EBRT should be accounted for by applying a margin around the tumor volume. Treatment preparation starts by delineating the gross palpable, visible or clinically demonstrable location of the tumor on a scan, yielding the gross tumor volume (GTV). Since microscopic disease spread surrounding the GTV is invisible on the scan, the GTV is expanded with a certain margin, resulting in the clinical target volume (CTV). An additional margin is applied to account for geometrical uncertainties in treatment planning (e.g., errors due to organ filling and movement) and delivery (e.g., set-up errors due to patient and beam positioning), which results in the planning target volume (PTV).

The PTV concept as described in the ICRU 62 report has been developed for EBRT, where the aim is to expand the dose distribution into a homogeneous plateau reaching beyond the CTV. For BT however, Tanderup et al. (2010) noted that: "a homogeneous dose cannot be obtained in and around a brachytherapy CTV", since adding PTV margins would lead to an undesirable dose escalation within the target. Applying a margin around the CTV to account for delineation uncertainties is thus not applicable for BT. Nevertheless, the margin recipe is currently used in BT practice.

1.1.2. Robust optimization So far, robust optimization and stochastic programming have only been applied to treatment planning models for EBRT. Stochastic programming considers the probability distribution of an uncertain parameter, for example through optimizing the expectation of the objective function or by restricting the probability of constraint violations (e.g. Chu et al. 2005, Olafsson and Wright 2006, Unkelbach and Ulfke

2004, Bohoslavsky et al. 2013, Fredriksson 2013). This inherently requires knowledge or assumptions regarding the probability distribution of the uncertain parameter. However, such information is often not available, as is the case in our application. Worst-case robust optimization on the other hand assumes an uncertainty region or a scenario-set in which the uncertain parameter resides. A worst-case robust optimization model only considers solutions that are robust feasible, i.e., solutions that are feasible for all possible realizations of the parameters. Among these solutions, a worst-case robust minimization (maximization) problem selects the solution that minimizes (maximizes) the maximum (minimum) over all possible parameter values or scenarios of the objective function (Ben-Tal et al. 2009). As a result, worst-case robust optimization yields treatment plans that are more robust than plans obtained with a stochastic programming approach, i.e., they perform better in the worst case scenario (Fredriksson 2012). Robust optimization for EBRT has been considered by various groups and can be applied to treatment planning models at three different levels: one can require robustness per calculation point (e.g. Chan et al. 2006, Bortfeld et al. 2008, Liu et al. 2012), per objective and constraint (Chen et al. 2011) or for the complete model (Fredriksson et al. 2011, e.g.). For a detailed comparison of these three approaches, we refer to Fredriksson and Bokrantz (2014). In our opinion, each constraint should be satisfied in all of the scenarios, and the robustness of each objective should be considered separately (as opposed to an aggregate of the objectives). In our model, where we have a single objective and several planning constraints, we apply robust optimization to each individual objective and constraint.

The methods described above have been applied to EBRT planning models for various types of uncertainties (e.g. organ motion or setup uncertainties), but none have considered delineation uncertainties. However, uncertainties in target volume delineation are known to be among the major causes of geometrical uncertainties (Weiss and Hess 2003) in EBRT and BT. The importance of accounting for any type of uncertainty in BT planning models has been emphasized by Kirisits et al. (2014). In particular, they note that intra- and inter-observer delineation variabilities, together with intra- and inter-fraction set-up uncertainties, contribute most to dosimetric uncertainty. Also Rylander et al. (2017) found that delineation uncertainties can lead to a degradation of dose. So far, methods for robust treatment planning only considered uncertainties that yield uncertainty in the location of the calculation point relative to the radiation source (e.g. through organ motion or setup uncertainties), while the structure the calculation point belongs to is fixed. This implies uncertainty in the dose rate (i.e., the dose per unit time), which is an important input parameter in treatment planning optimization models. On the other hand, delineation uncertainties do not change the location of a calculation point, so the dose rate remains fixed. Instead, they yield uncertainty in the structure a calculation point belongs to. The literature on robust optimization only considers uncertainty in input parameters. As our problem concerns uncertainty in index sets, we cannot use previously developed models, and a new approach is required.

1.2. Aim and contribution of the paper

The goal of this work is to take delineation uncertainties into account in the treatment planning optimization process for prostate HDR-BT using a worst-case robust optimization approach. For this, we need to develop a worst-case robust optimization method to incorporate uncertainties in index sets. We aim to reduce the risk of underdosing the CTV while still respecting the pre-defined OAR constraints. We compare our method to the classical margin approach.

The contributions of this work are (1) using computational methods to incorporate delineation uncertainties in the treatment plan optimization, (2) extending the robust optimization approach to account for uncertainties in index sets, (3) using robust optimization to deal with uncertainties in BT and (4) providing a speed-up for the nominal treatment planning model for BT that optimizes clinical objectives (Gorissen et al. 2013).

This paper is organized as follows. In Section 2, HDR-BT plan optimization is further clarified (Section 2.1) and the nominal model for HDR-BT of prostate cancer is introduced (Section 2.2). In Section 2.3 three methods for dealing with delineation uncertainties are discussed. An extension of the model to account for rectum delineation uncertainties is presented in Section 2.4, and methods for reducing the calculation times are presented in Section 2.5. The computational experiments and their results, along with a comparison of the three approaches, are presented in Section 3. A discussion and conclusion are provided in Sections 4 and 5, respectively.

2. Treatment plan optimization model

2.1. Dose prescription and plan evaluation

The dosimetric quality of a treatment plan is usually evaluated using dose-volume histogram (DVH) metrics. These metrics are denoted by $D_{x\%}(S)$ and $D_{x\text{cc}}(S)$, which reflect the minimum dose received by the hottest $x\%$ and x cc of the structure volume S, respectively, or by $V_{y\%}(S)$, which denotes the fraction of the structure volume S receiving at least $y\%$ of the prescribed tumor dose. A treatment plan should satisfy pre-set constraints on the DVH metrics. An example of a dose prescription protocol for prostate HDR-BT, for which the rectum and urethra are the most relevant OARs, is presented in Table 1. Here, $D_{90\%}(\text{PTV})$ is given as a percentage of the prescribed tumor dose. Note that, since in clinical practice uncertainties are accounted for using a CTV-to-PTV margin, dose prescriptions are defined for the PTV, not for the CTV. However, as in clinical practice the same prescriptions are used for the PTV, we will use them to assess plan quality of our methods as well.

Table 1 **Dose-volume criteria based on the protocol by Hoskin et al. (2007).**

PTV	Rectum	Urethra
$D_{90\%} \geq 100\%$	$D_{10\%} \leq 7.2$ Gy	$D_{10\%} \leq 10$ Gy
$V_{100\%} \geq 95\%$	$D_{2\text{cc}} \leq 6.7$ Gy	$D_{0.1\text{cc}} \leq 10$ Gy
$V_{150\%} \leq 55\%$	$D_{\max} \leq 8$ Gy	$D_{\max} \leq 10.6$ Gy
$V_{200\%} \leq 20\%$		

2.2. Nominal treatment planning model

The optimization models that are currently employed by treatment planning systems assign a penalty to each calculation point based on the difference between the planned and the prescribed dose, and minimize the total penalty (e.g., Lessard and Pouliot 2001). Since such penalties are a surrogate for the actual planning goals, which is to satisfy the pre-set DVH criteria, recently developed methods directly optimize the DVH metrics (Siauw et al. 2011, Gorissen et al. 2013, Holm et al. 2013). As a DVH-based optimization model uses clinically relevant objectives, we use the model from Gorissen et al. (2013).

In Gorissen et al. (2013), both catheter locations and dwell time distributions are optimized, where active dwell locations were given for each candidate catheter location with a

3 mm separation. The goal is to maximize the fraction of the target volume receiving at least the prescribed dose, denoted by $V_{100\%}(\text{CTV})$. The model is referred to as the linear dose-volume, or (LDV), model:

$$
\begin{aligned}
(LDV) \quad \max \quad & \frac{1}{|I_C|}\sum_{i\in I_C} v_i & & \\
\text{s.t.} \quad & \dot{d}_i^T t \geq v_i D_{pres} & \forall i \in I_C \quad & (1)\\
& \dot{d}_i^T t \leq L_R + (U_R - L_R)(1-w_i) & \forall i \in I_R \quad & (2)\\
& \sum_{i\in I_R} w_i \geq \tau_R |I_R| & & (3)\\
& t \geq 0 & & \\
& v_i \in \{0,1\} & \forall i \in I_C & \\
& w_i \in \{0,1\} & \forall i \in I_R & \\
& \text{[see Appendix for additional constraints (10) up to (14)]} & &
\end{aligned}
$$

Here, $\dot{d}_i \in \mathbb{R}_+^{|J|}$ denotes the vector with dose rates from each dwell position to calculation point i, I_C denotes the set of dose calculation points in the CTV, and J denotes the set of dwell positions. The optimization variable $t \in \mathbb{R}_+^{|J|}$ contains the (nonnegative) dwell times of all dwell positions. As a result, $\dot{d}_i^T t$ gives the total dose planned to be delivered to calculation point i. D_{pres} denotes the prescribed dose to the PTV, and v_i is an auxiliary variable that is equal to one if calculation point i receives at least the prescribed dose, and zero otherwise. This is enforced by the first constraint and the objective. The latter maximizes the fraction of calculation points receiving at least the prescribed dose. Constraint (2) ensures that no calculation point in the rectum receives a dose above U_R. The variable w_i equals 1 if calculation point i receives a dose below L_R, and 0 otherwise. Constraints (2) and (3) together restrict the fraction of calculation points in the rectum receiving a dose above L_R to be at most τ_R.

Additional constraints are included to restrict the dose to the urethra and to choose the number of catheters and their locations (see Appendix). These constraints are not provided in detail here, as they remain the same in the robust model.

2.3. Accounting for delineation uncertainties of the PTV

In this work we compare three approaches to incorporate delineation uncertainties in the plan optimization: an isotropic CTV-to-PTV margin as is currently used in the clinic, a scenario-based CTV-to-PTV margin, and a robust optimization approach. Each of these methods is explained below.

2.3.1. Isotropic CTV-to-PTV margin Despite the fact that a CTV-to-PTV margin seems unsuitable for HDR BT, it is currently used in the clinic to account for delineation uncertainties. As we compare our robust approach to the current clinical standard, we consider the isotropic CTV-to-PTV margin here. For model (LDV), this implies that the set I_C is replaced by I_P, the latter of which is the set of calculation points in the PTV. An isotropic margin of 2 mm is used.

2.3.2. Scenario-based CTV-to-PTV margin Given a nominal delineation of the CTV and using measurements on delineation inaccuracies from Smith et al. (2007), we generate scenarios for possible delineations and the corresponding sets of CTV calculation points. Since the CTV-to-PTV margin is smaller than the delineation uncertainty reported in the literature, the PTV does not fully contain all the CTV scenarios, and I_P does not contain all of the calculation points that are in the CTV according to at least one scenario. Therefore, we also test the use of a scenario-based margin. The union of all CTV scenarios is considered as the PTV, and the set of all calculation points that may be in the CTV according to our scenario set, denoted by $\tilde{I}_C$, is used instead of the set I_C in (LDV).

2.3.3. Robust optimization We use a scenario-based approach for our robust optimization model. For each scenario in the set S, we know the set of calculation points within the CTV. This information is stored in matrix C, a $|S| \times |I|$ matrix, where each row corresponds to a scenario and each column corresponds to a calculation point. The entry on the s^{th} row in the i^{th} column equals 1 if calculation point i resides in the CTV for scenario s, and zero otherwise. This matrix is used to calculate CTV coverage for each scenario s as $V^s_{100\%}(\text{CTV}) = C_s v / C_s e$, where C_s denotes row s of matrix C, e is the all-ones vector and v is as before. The numerator counts the number of calculation points that receive at least the prescribed dose and are in the CTV according to scenario s, while the denominator counts the number of calculation points in the CTV according to scenario s.

The robust counterpart of the (LDV) model is formulated as:

$$
\begin{aligned}
(RC) \quad \max \quad & V & & & \\
\text{s.t.} \quad & V \leq \frac{C_s v}{C_s e} & \forall s \in S & & (4)\\
& \dot{d}_i^T t \geq D_{pres} v_i & \forall i \in \tilde{I}_C & & (5)\\
& t \geq 0 & & & \\
& v_i \in \{0,1\} & \forall i \in \tilde{I}_C & & (6)\\
& \dot{d}_i^T t \leq L_R + (U_R - L_R)(1 - w_i) & \forall i \in I_R & & \\
& \sum_{i \in I_R} w_i \geq \tau_R |I_R| & & & \\
& w_i \in \{0,1\} & \forall i \in I_R & & \\
& \text{[additional constraints (10) up to (14) in Appendix 5].} & & &
\end{aligned}
$$

Initial tests show that (RC) yields a risk of overdosage, reflected by $V_{200\%}(\text{CTV})$ exceeding the desired maximum level in Table 1. Therefore, we added the following constraints to (RC) in order to limit $V_{200\%}$ for each scenario:

$$
\begin{aligned}
& \dot{d}_i^T t \leq 2 D_{pres} + M u_i && \forall i \in \tilde{I}_C && \\
& \frac{C_s u}{C_s e} \leq 0.2 && \forall s \in S && (7)\\
& u_i \in \{0,1\} && \forall i \in \tilde{I}_C, &&
\end{aligned}
$$

where u_i is a binary variable that is 0 only if calculation point i receives at most twice the prescription dose and 1 otherwise, and M is an arbitrary large number. For scenario

s, we have $V^s_{200\%}(CTV) = C_s u / C_s e$, which is restricted to be at most 0.2 according to the protocol by Hoskin et al. (2007), see Table 1.

The size of (LDV) and (RC) can be found in Table 2, where we only consider the constraints and variables related to the target volume, and skip the constraints and variables corresponding to OAR sparing, maximum dwell times and catheter choice as these are identical for both models. This table clearly shows the major advantage of our approach: the number of binary variables does not increase with the number of scenarios $|S|$, but only with the number of calculation points in the uncertainty region $|\tilde{I}_C|$. The number of constraints increases linearly with $|S|$.

Table 2 Problem sizes for (LDV) and (RC).

	(LDV)	(RC)
Number of binary variables	$\|I_P\|$	$2\|\tilde{I}_C\|$
Number of continuous variables	$\|J\|$	$\|J\|$
Number of constraints	$\|I_P\|$	$2\|S\| + 2\|\tilde{I}_C\|$

2.4. Accounting for uncertainties in rectum delineation

Delineation uncertainties in the OARs are not accounted for in current clinical practice. The delineation of the urethra is indeed rather accurate, as a Foley catheter is usually inserted that is well visible on images acquired prior to treatment planning. There is however variability in the rectal wall thickness. Note that, when using transrectal ultrasound imaging, the position of the posterior rectal wall is invisible and thus not delineated. This however is irrelevant, as this volume receives little or no dose. Delineation uncertainties in the cranial and caudal as well as the lateral direction of the rectum are irrelevant, as these areas receive hardly any dose. Therefore, we only consider variabilities in delineating the anterior rectal wall.

Variabilities in delineation of the rectum can be accounted for using a margin or following a robust optimization approach. We consider the scenario-based margin.

For the robust optimization approach, we need to derive the robust counterpart of constraints (2) and (3). Following the same approach as for the PTV, we consider a set of scenarios for the delineation of the rectum. The rows of the matrix R correspond to the calculation points, and the columns correspond to the scenarios. An element R_{is} equals 1 if calculation point i is in the rectum according to scenario s, and 0 otherwise. We obtain the following robust counterpart of constraints (2) and (3):

$$\begin{aligned} &\dot{d}_i^T t \leq L_R + (U_R - L_R)(1 - w_i) && \forall i \in \tilde{I}_R \\ &\frac{R_s^T w}{R_s^T e} \geq \tau_R, \end{aligned} \tag{8}$$

where $\tilde{I}_R$ is the set containing all calculation points that are in the rectum according to at least one scenario.

It could be the case that the sets $\tilde{I}_C$ and $\tilde{I}_R$ overlap in space. However, as a different set of calculation points is generated for each structure, it is not possible that one calculation point belongs to two structures for some scenarios, and hence that two different restrictions are imposed for a single calculation point. This allows us to consider the worst case for the objective function and each constraint separately.

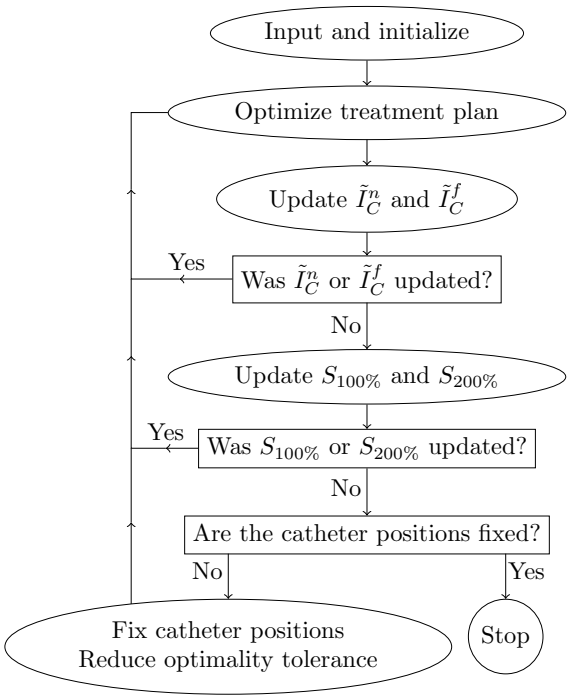

Figure 3 **Flow chart of the algorithm used for solving** (RC)**. Details of the procedures "Input and initialize", "Update** $\tilde{I}_C^n$ **and** $\tilde{I}_C^f$**" and "Update** $S_{100\%}$ **and** $S_{200\%}$**" can be found in Appendix B.**

2.5. Reduction of solution times

Even though the number of binary variables increases only with the size of the uncertainty region, the solution time becomes more than several days, which is too long for the model to be used in clinical practice, or may not be solved at all due to memory issues. Therefore, we propose several heuristics that are combined into one algorithm to speed up the optimization. The algorithm is summarized in Figure 3.

First of all, for the CTV we relax the requirement $v_i \in \{0,1\}$ to $v_i \in [0,1]$. Besides strongly reducing the number of binary variables and thus the calculation times, this relaxation has an appealing interpretation. Both for the binary and the continuous case, we have that $v_i = 1$ when calculation point i receives at least the prescribed dose. When i receives a dose below the prescription dose, the binary v_i equals 0, whereas the continuous variable v_i equals the delivered dose as a fraction of the prescribed dose. Thus, when using continuous variables, the dose to calculation point i is still pushed upwards, even when calculation point i does not receive the prescribed dose.

Secondly, optimization speed is improved by exploiting the fact that calculation points in close proximity to the catheters are likely to receive a high dose, while calculation points at a larger distance are likely to receive a low dose. This is due to the inverse quadratic relation between the dose rate from a dwell position to a calculation point and the distance between the two points (Nath et al. 1995). It may thus not be necessary to optimize for all calculation points in $\tilde{I}_C$. Therefore, we initially assume that the calculation points near the catheters, denoted by the set $\tilde{I}_C^n$, receive at least the prescribed dose (i.e., $v_i = 1 \ \forall i \in \tilde{I}_C^n$), and that calculation points far from the catheters, denoted by $\tilde{I}_C^f$, do not receive more than $2D_{pres}$ (i.e., $u_i = 1 \ \forall i \in \tilde{I}_C^f$). An example of such sets is illustrated in Figure 4. One could define these sets based on e.g. the distance to the nearest dwell position. We define these sets based on the outcome of the nominal optimization model:

all calculation points that receive a dose above D_{pres} or below $2D_{pres}$ when applying the nominal treatment plan optimized with (LDV) are included in $\tilde{I}_C^n$ and $\tilde{I}_C^f$, respectively. Note that for the calculation points outside $\tilde{I}_C^n$ and $\tilde{I}_C^f$ no assumptions on the received dose levels are made. (RC) is optimized where constraints (4) and (7) only apply to calculation points in $\tilde{I}_C \setminus \tilde{I}_C^n$ and $\tilde{I}_C \setminus \tilde{I}_C^f$, respectively. After optimization, we check for each of the calculation points in $\tilde{I}_C^n$ and $\tilde{I}_C^f$ whether they indeed receive a sufficiently high and low dose as respectively was assumed. If not, we exclude a predetermined number of the coldest and hottest calculation points from the sets $\tilde{I}_C^n$ and $\tilde{I}_C^f$, respectively, and re-optimize using the previous optimal treatment plan as a starting solution. In our experiments, we excluded $\eta_{100\%}|\tilde{I}_C|$ and $\eta_{200\%}|\tilde{I}_C|$ calculation points per iteration from the sets $\tilde{I}_C^n$ and $\tilde{I}_C^f$, respectively, where $\eta_{100\%},\ \eta_{200\%} \in [0,1]$. This process is continued until the number of calculation points for which we made an incorrect assumption is sufficiently low (see Figure 3 and Appendix B).

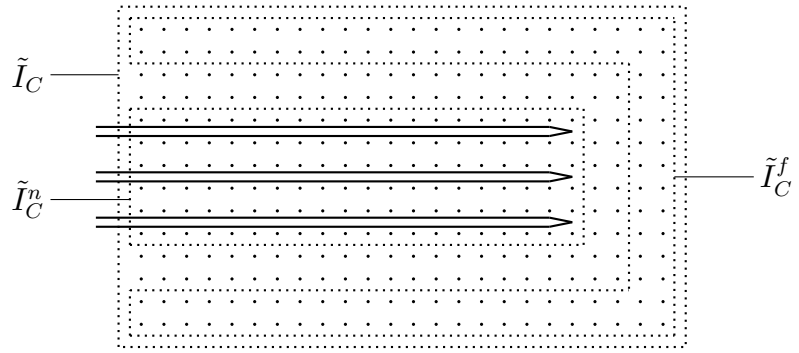

Figure 4 **Example of sets of calculation points that are likely to receive a high dose ($\tilde{I}_C^n$) or a low dose ($\tilde{I}_C^f$).**

Thirdly, we implement the adversarial approach (Bienstock and Özbay 2008), which implies that only a subset of the scenarios is included in the optimization. We define two sets, $S_{100\%} \subset S$ and $S_{200\%} \subset S$. The first, $S_{100\%}$, contains those scenarios that are accounted for in the model when determining the worst case $V_{100\%}$ in the objective, while the second, $S_{200\%}$, contains those scenarios for which we require $V_{200\%} \leq 0.2$. The robust model is thus solved with S replaced by $S_{100\%}$ in the $V_{100\%}$ objective and S replaced by $S_{200\%}$ for the constraint on $V_{200\%}$ of the CTV. This means that the resulting optimal solution is robust with respect to the scenarios in $S_{100\%}$, and $V_{200\%} \leq 0.2$ for the scenarios in $S_{200\%}$. Using the optimal treatment plan, the scenario with the lowest $V_{100\%}$(CTV) is added to $S_{100\%}$, unless it is in the set already. Furthermore, if the highest $V_{200\%}$ over all scenarios is larger than 0.2, the corresponding scenario is added to $S_{200\%}$. The new model instance is solved using the previous optimal solution as a starting point. The scenario sets are updated until no more scenarios are added to either of the sets. Initially, $S_{100\%}$ only contains the nominal scenario and $S_{200\%}$ is empty (see Figure 3 and Appendix B).

The final iterations in the optimization are often spent on improving the optimality bound while the objective value is hardly or not at all improved. Furthermore, improving the objective a little bit may in practice have little effect on the dose to the tumor or OARs. It is thus not necessary to solve the model to optimality. This particularly holds for optimization of the catheter configuration: the minor improvements in the final iterations are likely to result from a change in dwell times, not from a change in catheter configuration. Therefore, we optimize the robust model up to a pre-determined optimality gap g_1, fix the

catheter configuration, and continue optimizing the beam-on times up to a second, smaller optimality gap g_2 (see Figure 3).

3. Computational experiments

3.1. Patient and source data

Data from six prostate cancer patients were used to test our robust optimization method. Delineations of the CTV, rectum and urethra were obtained from the treatment planning system, see Table 3 for the structure volumes. The data from patients 1, 2 and 3 were obtained from the planning system HDRplus (Eckert and Ziegler BEBIG GmbH, Berlin, Germany), and have previously been used in Gorissen et al. (2013) and in Balvert et al. (2015). The delineations were originally made on images obtained with transrectal ultrasound. For these three patients, catheter- and dwell locations could be obtained from the treatment planning system as well. The data from patients 4, 5 and 6 were obtained from Oncentra Brachy (Nucletron, Veenendaal, The Netherlands), and were previously used in Deist and Gorissen (2016). The images of these patients were made using computed tomography. No data on catheter locations and dwell positions were available from the planning system, and were therefore added by virtually placing a template in front of the perineum using MATLAB Release 2012b (The Mathworks, Inc., Natick, USA).

For all six patients, calculation points were hexagonally distributed over the structures using MATLAB Release 2012b (The Mathworks, Inc., Natick, USA), of which details can be found in Table 4. We used the same number of dose calculation points for optimization as the treatment planning system did, and used a larger set for the dosimetric evaluation of the treatment plan in order to obtain more accurate DVH measures. As the first three patients were imaged with transrectal ultrasound, only the anterior rectum wall was delineated, while for patients 4, 5 and 6 the complete rectum volume was delineated on the computed tomography scans. As the density of the calculation points was the same for all patients, the set of rectum calculation points is larger for patients 4, 5 and 6 than for the first three patients.

Dose rates were calculated according to the TG-43 formalism (Nath et al. 1995) for which ^{192}Ir source-specific parameters were obtained from Granero et al. (2006). The number of possible catheter positions and corresponding dwell positions can be found in Table 5.

Table 3 Tissue structure volumes (cc).

Structure	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6
Nominal CTV	31.7	55.2	47.6	31.0	40.1	32.7
Isotropic margin PTV	39.6	62.8	63.2	37.0	43.8	37.4
Scenario-based PTV	54.5	97.3	70.8	57.4	57.6	59.9
Nominal rectum	6.8	7.5	9.4	19.5	20.7	16.3
Scenario-based rectum	8.5	8.8	12.4	20.3	20.2	16.5
Urethra	2.1	2.6	2.3	1.3	3.0	1.3

Table 4 Number of dose calculation points. Pt=patient, cps=calculation points.

	Optimization					
Structure (set of cps)	Pt 1	Pt 2	Pt 3	Pt 4	Pt 5	Pt 6
Nominal CTV (I_C)	1750	1759	1743	1757	1749	1743
Isotropic margin PTV (I_P)	1777	1754	1757	1790	1772	1795
Scenario-based PTV ($\tilde{I}_C$)	2959	2766	2729	3322	3228	3109
Nominal rectum (I_R)	249	253	251	257	259	256
Scenario-based rectum ($\tilde{I}_R$)	312	298	333	268	265	259
Urethra (I_U)	465	489	488	479	482	465
	Evaluation					
Nominal CTV (I_C)	8108	8112	8246	8676	8166	8058
Isotropic margin PTV (I_P)	7955	7892	8423	8214	8273	8133
Scenario-based PTV ($\tilde{I}_C$)	20790	23393	26654	15951	14441	15365
Nominal rectum (I_R)	2586	2533	2542	6252	6283	6124
Scenario-based rectum ($\tilde{I}_R$)	3205	3359	3450	6552	6441	6625
Urethra (I_U)	2114	2163	2314	1992	2090	1956

Table 5 Number of catheters and dwell potitions. Pt=patient, cps=calculation points.

	Pt 1	Pt 2	Pt 3	Pt 4	Pt 5	Pt 6
Catheters	40	49	43	51	51	71
Dwell positions, original set	369	711	716	414	454	481
Dwell positions, extended set	432	813	836	657	582	653

3.2. Experiment setup

To define the uncertainty region for the CTV, we used scenarios obtained from the original contours by stretching or shrinking the delineated target volume in the left, right, anterior, posterior, superior and inferior direction. The centroid-to-surface distances in each direction were varied independently of each other. Distances were assumed to vary at most two standard deviations from the mean (delineated) distance. Standard deviations were obtained from Smith et al. (2007). For patients 1, 2 and 3, with whom transrectal ultrasound imaging was used, the standard deviations were 2.2 mm for the superior and inferior directions, and 1.15 mm for the remaining directions. Patients 4, 5 and 6 were imaged using computed tomography, and their standard deviations were 1.15 mm in the cranial and caudal directions, 1.7 mm in the lateral directions, and 1.6 mm and 2 mm in the anterior and posterior directions, respectively. For patient 2, the original, smallest and largest possible CTV shapes are depicted in Figure 5.

An uncertainty region for the rectum delineation was constructed through scenario generation as well. However, for the rectum only variations in the anterior direction are of interest, as this is where the highest dose is deposited. As a result, only three scenarios are included in the optimization: the smallest, the nominal and the largest possible delineation. The smallest (largest) scenario was generated by shrinking (stretching) the rectum volume in the anterior direction by two standard deviations. We assume a standard deviation of 0.5 mm.

This method of scenario generation is rather straightforward as it neglects the fact that delineations of an observer may deviate from those delineated by others in a consistent manner, for example one observer may always draw larger shapes than his/her colleague. The assumption of independent deviations in each direction may thus be invalid. However, to the best of our knowledge, there is no data available on this dependency. Our method can be easily adapted if such data would become available.

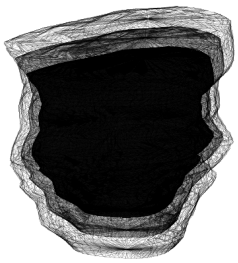

Figure 5 **Vectorized figure of minimal (dark gray), delineated (medium gray) and maximum (light gray) CTV shape.**

In order to limit the number of scenarios taken into account in the optimization process, we considered only the minimum, nominal and maximum distance for each direction, which gives $3^6 = 729$ possible shapes. Note that the assumption of independent deviations for each direction may result in unrealistically large or small shapes, e.g., when we fully stretch the shape in each direction, the volume becomes clinically unrealistically large. Therefore, all scenarios with a clinically unrealistic CTV volume (smaller than 20 cc or larger than 65 cc) were excluded before optimization. This results in 596, 534, 602, 666, 729 and 729 scenarios for the six respective patients. In order to evaluate the plan quality, over 5,000 CTV scenarios were generated for each patient by randomly drawing centroid-to-surface distances for each direction. Again, extremely small and large scenarios were excluded. Recall that for the rectum only the uncertainty in the anterior rectal wall is relevant, which allows us to generate scenarios by stretching or shrinking the rectum in the anterior direction only. We generated 101 scenarios, for which the factors by which the rectum was shrunk or stretched were sampled at equal distances on the interval from the minimum to the maximum possible deviation.

Models were compared based on DVH evaluation criteria for all scenarios as well as solution times. The models were solved using the Gurobi 5.5 optimizer (Gurobi Optimization, Inc., Houston, USA) interfaced with MATLAB Release 2012b on a computer with an Intel i7-2670 QM processor.

The dwell positions that correspond to each catheter location are predefined by the planning software. In our test cases, only the dwell positions that are in the PTV were activated and thus included in the optimization process. As the union of the CTV scenarios is larger than the original target volume, there are dwell positions that were not included in the optimization process whereas they may be inside the target volume. It may thus be necessary to include additional dwell positions for each catheter. This is illustrated in Figure 6, where the PTV is delineated in white. The red dwell positions are included in the original dataset, which can be extended by including the white dwell positions as well. All models are therefore solved once using the original and once using the extended set of dwell positions.

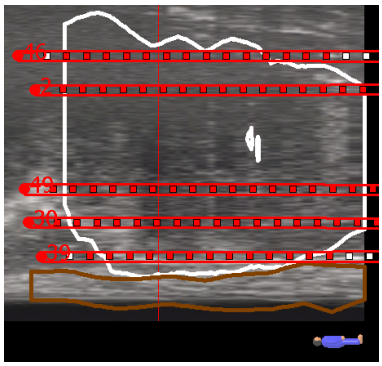

Figure 6 **Sagittal transrectal ultrasound scan with the PTV delineated in white. The red dwell positions are included in the original dataset, the white ones are added to obtain the extended set of dwell positions.**

Table 6 **Overview of the settings used for each optimization approach.**

	Method to account for uncertainties in CTV	in rectum	Dwell positions
INO	Isotropic margin	None	Original set
RNO	Worst case robust	None	Original set
RRO	Worst case robust	Worst case robust	Original set
SNO	Scenario-based margin	None	Original set
SSO	Scenario-based margin	Scenario-based margin	Original set
INE	Isotropic margin	None	Extended set
RNE	Worst case robust	None	Extended set
RRE	Worst case robust	Worst case robust	Extended set
SNE	Scenario-based margin	None	Extended set
SSE	Scenario-based margin	Scenario-based margin	Extended set

For each patient a total of ten treatment plans are optimized. The different optimization approaches are denoted by a three-letter code. The first letter denotes the approach taken to account for delineation uncertainties of the CTV, where I denotes the use of an isotropic CTV-to-PTV margin, S denotes the scenario-based margin and the robust optimization approach is denoted by R. The second letter indicates the approach followed to account for delineation uncertainties of the rectum, where N denotes the use of the nominal rectum delineation, and S and R are as before. Finally, the third letter indicates whether the original (O) or the extended (E) set of dwell positions is used. For example, INO denotes the optimization approach followed in the clinic, where an isotropic CTV-to-PTV margin is employed, the nominal delineation is included for the rectum (so rectum delineation uncertainty is ignored) and the original set of dwell positions is used. An overview of the various settings used for the five optimization approaches can be found in Table 6.

The optimality gap g_2 is chosen as the smallest value for which the algorithm terminates within 15 minutes (a clinically acceptable optimization time) when optimizing RNO. This value differs per patient, and can only be found through a trial-and-error search. The optimality gaps used can be found in Table 7. Note that for patient 2 the optimality gap had to be adapted for the margin approach, as the optimizer ran out of memory when using a smaller value for g_2. The parameter g_1 is set to be twice g_2. Parameters $\eta_{100\%}$ and

$\eta_{200\%}$, the fractions of calculation points added to the sets $\tilde{I}_C^n$ and $\tilde{I}_C^f$ in each iteration, respectively, were set at 0.10.

Table 7 Optimality gaps (g_2 in the algorithm) used for the individual patients. The robust approach comprises models RNO, RRO, RNE and RRE, and the scenario-based margin approach comprises models SNO, SSO, SNE and SSE.

	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6
Gap (%) robust models	0.400	0.150	0.025	0.400	0.100	0.200
Gap (%) margin models	0.400	1.000	0.025	0.400	0.100	0.200

3.3. Numerical results

In this section, we first look at the effects of using continuous instead of binary variables v by comparing the results obtained with (LDV) to results obtained with its relaxation (Section 3.3.1). Next, in Section 3.3.2 we compare the proposed methods to account for delineation uncertainties. Finally, we look at the effects of changing the algorithm parameters in Section 3.3.3.

3.3.1. Binary versus continuous variables v In order to assess the effects of using binary instead of continuous variables, we first consider the (LDV) model optimized with and evaluated in the nominal scenario only (INO). The relaxation of the (LDV) model for the nominal case results in a large reduction of solution times, without a significant compromise in plan quality (Table 8). Note that for patient 1 the dose requirement $D_{0.1\text{cc}}$ is slightly violated, though the violation is very small and thus clinically irrelevant. The solution times of the robust model without the relaxation of v_i and the proposed algorithm from Figure 3 are unacceptable: after 5 hours, the optimality gap is still approximately 5%, 15% and 10% for patients 1, 2 and 3, respectively. This makes the model with binary variables unusable and the relaxation necessary.

3.3.2. Comparison of the treatment plan optimization approaches Treatment plans were generated for each patient using the ten models from Table 6, where the solution times of the margin and the robust models were reduced using the algorithm in Figure 3. The nominal model was solved to optimality without using the algorithm.

Treatment plan quality in terms of DVH parameters We are interested in the distribution of the CTV and rectum DVH parameters over the scenarios. Therefore, in Figure 7, we show the cumulative distribution of $D_{90\%}$(CTV), $V_{100\%}$(CTV), $V_{150\%}$(CTV), $V_{200\%}$(CTV), $D_{10\%}$(Rectum) and $D_{2\text{cc}}$(Rectum) for patient 1. Similar figures for patients 2 up to 6 are included in the appendix (Figures 8 up to 12). For example, in Figure 7(c) we see that for the plan generated for patient 1 using the INO approach (solid black curve), approximately 20% of the scenarios had a $V_{150\%}$(CTV) of 0.45 or lower. The gray area denotes values below the desired minimum value for $D_{90\%}$(PTV) and $V_{100\%}$(PTV), and above the desired maximum value for the remaining DVH parameters. For completeness, Table 11 in Appendix C.2 shows the DVH parameters for the urethra. This table shows that the constraint $D_{10\%}$(Urethra) is always satisfied.

Recall that our aim is to achieve high $D_{90\%}$(CTV) and $V_{100\%}$(CTV), where we aim to achieve minimum levels of 1.00 and 0.90, respectively. A curve is thus superior to another

Table 8 Comparison of treatment plans generated with (LDV) and plans generated with a relaxation of (LDV).

DVH parameter	unit	Patient 1		Patient 2		Patient 3	
		(LDV)	Relaxation	(LDV)	Relaxation	(LDV)	Relaxation
$D_{90\%}(\text{PTV}) \geq 1.00$	%	1.13	1.15	1.07	1.07	1.05	1.05
$V_{100\%}(\text{PTV}) \geq 0.90$	%	0.99	1.00	0.97	0.97	0.97	0.97
$V_{150\%}(\text{PTV}) \leq 0.55$	%	0.50	0.51	0.25	0.29	0.26	0.27
$V_{200\%}(\text{PTV}) \leq 0.20$	%	0.20	0.26	0.09	0.13	0.10	0.11
$D_{10\%}(\text{Rectum}) \leq 7.2$	Gy	6.8	7.3	7.2	7.2	6.9	7.0
$D_{2\text{cc}\%}(\text{Rectum}) \leq 6.7$	Gy	5.8	6.3	6.2	6.2	6.0	6.1
$D_{10\%}(\text{Urethra}) \leq 10$	Gy	9.8	10.0	9.8	9.7	9.5	9.7
$D_{0.1\text{cc}\%}(\text{Urethra}) \leq 10$	Gy	10.0	10.2	10.0	9.9	9.7	9.8
Solution time	s	14	8	385	26	25	14
DVH parameter	unit	Patient 4		Patient 5		Patient 6	
		(LDV)	Relaxation	(LDV)	Relaxation	(LDV)	Relaxation
$D_{90\%}(\text{PTV}) \geq 1.00$	%	1.13	1.13	1.12	1.14	1.14	1.16
$V_{100\%}(\text{PTV}) \geq 0.90$	%	0.98	0.98	0.98	0.98	0.99	0.99
$V_{150\%}(\text{PTV}) \leq 0.55$	%	0.47	0.49	0.42	0.45	0.50	0.50
$V_{200\%}(\text{PTV}) \leq 0.20$	%	0.26	0.29	0.19	0.22	0.26	0.24
$D_{10\%}(\text{Rectum}) \leq 7.2$	Gy	5.5	5.6	5.2	5.2	5.6	5.8
$D_{2\text{cc}\%}(\text{Rectum}) \leq 6.7$	Gy	5.5	5.6	5.2	5.2	5.3	5.5
$D_{10\%}(\text{Urethra}) \leq 10$	Gy	9.8	10.0	9.8	9.8	9.7	9.9
$D_{0.1\text{cc}\%}(\text{Urethra}) \leq 10$	Gy	9.7	9.9	9.8	9.8	6.5	6.4
Solution time	s	2302	498	116	10	35	9

curve if it lies further to the right, and we'd preferably see a curve that lies completely to the right of the minimum level. Similarly, for $V_{150\%}(\text{CTV})$, $V_{200\%}(\text{CTV})$, $D_{10\%}(\text{Rectum})$ and $D_{2\text{cc}}(\text{Rectum})$ curves that lie further to the left are superior, as these parameters are desired to be small.

First we consider the models where rectum uncertainties are not accounted for. When comparing RNO and RNE to INO and INE, respectively, we observe a shift of the complete curve of $V_{100\%}(\text{CTV})$ towards the higher values in almost all cases. Exceptions are RNO for patient 2, which performs better than INO for the worst 15% but worse for the remaining 85% of the simulated delineations, RNE for patient 6, which performs better than INE for the worst 10% but worse for the remaining 90% of the simulations, and RNE for patient 1, where the curve lies slightly more towards the lower values than the curve corresponding to INE. Except for patient 2, SNO and SNE always yield better $V_{100\%}(\text{CTV})$ than INO and INE, respectively. The scenario-based margin and the robust optimization approach perform equally well in terms of target coverage.

Including rectum delineation inaccuracies in the optimization process, i.e. using the RR and SS models instead of the RN and SN models, respectively, yields a lower rectum dose for patients 1, 2 and 3. For patients 2 and 3 this also implies a reduced target coverage. Including rectum delineation uncertainties does not affect treatment plans for patients 4, 5 and 6, as the rectum dose is already sufficiently low for these three patients: the constraint on the rectum dose was never active.

The use of a scenario-based CTV-to-PTV margin in the SN and SS models often yields an overdosage to the target volume in terms of $V_{200\%}(\text{CTV})$ compared to RN and RR. This becomes clear from the shift of the corresponding curves towards the right for patients 1, 4, 5 and 6. This shift is not visible for patients 2 and 3, though $V_{200\%}(\text{CTV})$ never exceeds the maximum allowed value of 0.2 for these two patients and thus never was an active constraint. The risk of an overdose is much lower for the RN and RR models.

When comparing the use of an extended set of dwell positions to using the original set for the same model (e.g. compare INO to INE or RRO to RRE), we see that in almost all cases the extended set allows for a reduction in $V_{150\%}(\text{CTV})$ and $V_{200\%}(\text{CTV})$ with little or no compromise on $V_{100\%}(\text{CTV})$. It does however often yield an increased rectum dose.

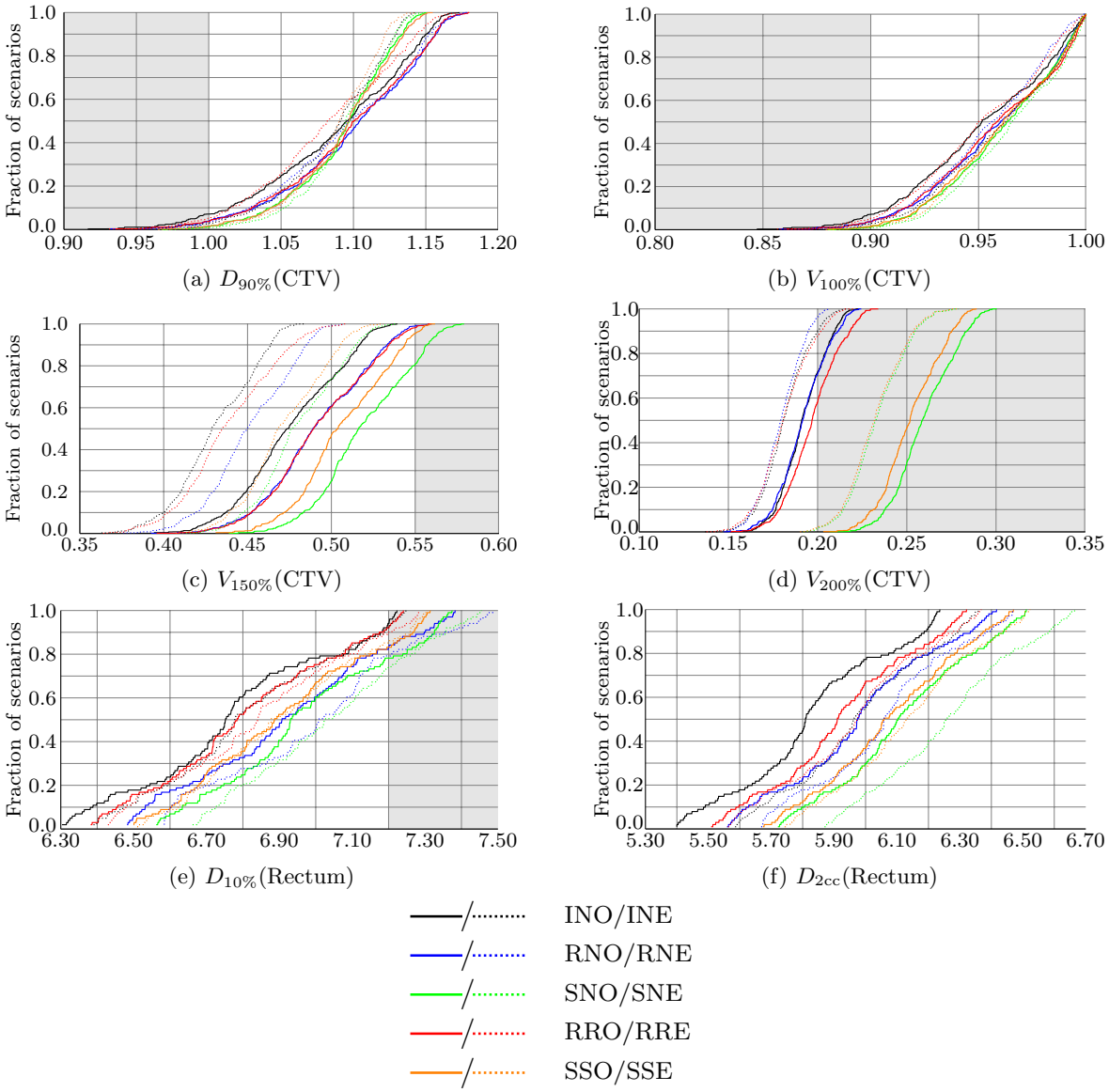

Figure 7 **DVH metrics for patient 1 obtained with various models, evaluated in the evaluation dataset. The gray area indicates scenarios where the DVH requirement from Table 1 was not met.**

Solution times The solution times of all methods for all patients are reported in Table 9. The solution times for RN, RR, SN and SS approaches are still clinically acceptable due to the choice of the accepted optimality gaps. Note that the solution times of RR and SS, i.e., those models where rectum delineations are accounted for, are always the same.

Table 9 Solution times.

	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6
INO	42	385	25	2302	116	35
RNO	571	98	183	912	220	528
RRO	65	313	256	344	493	496
SNO	880	123	213	612	844	362
SSO	63	313	260	335	481	496
INE	20	439	31	804	83	12
RNE	325	218	351	53	75	150
RRE	41	88	101	75	348	365
SNE	399	87	324	52	83	144
SSE	41	87	101	74	347	366

3.3.3. Changing the algorithm's input parameters Several parameters that form an input to the optimization algorithm can affect the algorithm's performance in terms of both objective value/plan quality and solution times. In this section, we look at the effects of changing the accepted optimality gaps and the sizes of the volume fractions to be removed from the sets of calculation points.

Optimality gaps g_1 and g_2
In each iteration, the algorithm optimizes an instance until the optimality gap g_1 is reached. After all stopping criteria have been satisfied (see Figure 3), the catheter positions are fixed and the dwell times are further optimized where in each iteration an instance is optimized up to optimality gap g_2, which is smaller than g_1.

In order to see the effects of choosing a different optimality gap, we optimized RNO for various values of g_2 for patients 1 and 3. We have evaluated the optimizations in terms of solution time, objective function value and worst case $V_{100\%}$(CTV) for the set of scenarios included in the optimization as well as the simulated scenarios that were used in our assessment of plan quality before. Detailed results are shown in Table 12 in Appendix D. We did not consider gaps smaller than 0.025%, as this is very small already. Optimizing RNO for patient 1 for $g_2 = 0.2\%$ took more than three hours, so the model was not optimized for gaps smaller than 0.3%. For patient 1, optimizing up to a g_2 smaller than 0.4% yields clinically unacceptable solution times. For patient 3 on the other hand, setting $g_2 = 0.4\%$ gives a much lower worst case $V_{100\%}$(CTV) than with $g_2 = 0.025\%$: 87.8 versus 92.7. It is thus necessary to determine a suitable g_2 for each patient individually.

Fractions of calculation points $\eta_{100\%}$ and $\eta_{200\%}$
In each iteration, the model is optimized for subsets of the calculation points. After each iteration, the algorithm checks whether the worst case $V_{100\%}$(CTV) and $V_{200\%}$(CTV) obtained for the optimized instance are sufficiently close to the real values. If this is not

the case, the sets of calculation points are extended. Recall that the parameters $\eta_{100\%}$ and $\eta_{200\%}$ determine the fractions of $\tilde{I}_C$ that are removed from $\tilde{I}_C^n$ and $\tilde{I}_C^f$, respectively. In order to see the effect of the choices for these parameter values, we have optimized RNO for patient 1 for various values of $\eta_{100\%}$ and $\eta_{200\%}$. The results presented in Table 13 in Appendix D show that the choice of $\eta_{100\%}$ and $\eta_{200\%}$ hardly influences the objective function value and the worst case $V_{100\%}(\text{CTV})$, but does significantly influence the solution times. There is however no clear relation between the choice of η and the solution times. This can be explained by the trade-off that one makes when choosing η: a small η results in adding few calculation points to the optimization problem in the next iteration, which makes the optimization fast but could result in the need for more iterations. On the other hand, a large η adds many calculation points to the optimization in each iteration, hence slowing down the optimization but reducing the number of iterations needed.

4. Discussion

In EBRT it is common practice to account for setup- and delineation uncertainties by applying a PTV margin around the CTV (International Commission on Radiation Units and Measurements 1999). However, a PTV margin approach is questionable for brachytherapy. First of all, often an isotropic margin is used, whereas the delineation uncertainties vary non-isotropically (Smith et al. 2007). Furthermore, Tanderup et al. (2010) argue that CTV-to-PTV margins cannot be applied to brachytherapy, since delivering a homogeneous target dose that reaches beyond the CTV would require a dose escalation in the interior of the CTV. These two observations were the motivation to develop our robust treatment planning method, and to compare it against the classical margin approach and against a margin approach that is based on scenarios of target volume delineations.

Our results indicate that target coverage is improved by using the scenario-based margin model instead of the isotropic 2 mm CTV-to-PTV margin, which implies that the 2 mm margin used in the clinic may be insufficient. This particularly holds for the anterior and posterior directions of the CTV, where the delineation uncertainty is larger than in other directions (Smith et al. 2007). We conclude that an isotropic margin is indeed not adequate. A downside of the scenario-based margin model is an increased $V_{200\%}(\text{CTV})$, which is in agreement with the findings from Tanderup et al. (2010). The robust treatment planning model results in an improvement in target coverage similar to the scenario-based margin approach. Additionally, in this model an overdosage, reflected in excessively high values for $V_{200\%}(\text{CTV})$, can be prevented by adding a constraint that requires $V_{200\%}(\text{CTV})$ to be below a preset level for each of the scenarios. This constraint does not work well for the margin approach, since in that case no individual scenarios are considered and thus the constraint on $V_{200\%}(\text{CTV})$ applies to an extended CTV only, which is a rather large volume. As a result, individual scenarios were not protected from overdosage.

The dose escalation in the CTV may be caused by the absence of dwell positions in the margin volume. Therefore, we optimized a treatment plan with the scenario-based margin approach using additional dwell positions located in the scenario-based margin. This did not improve the treatment plan quality, and we can conclude that dose escalation inherent to the margin approach cannot be prevented by including more dwell positions.

Uncertainties in the delineation of the rectum yield a risk of a rectal overdosage for some patients. Our results show that, for those patients, accounting for rectum delineation uncertainties reduces this risk. For two out of three patients for whom this risk was reduced, this comes at a cost of a reduced coverage of the target volume.

The solution time and the optimality gap both vary strongly among patients. Note that solution times vary for the nominal model as well. Furthermore, the solution time of the robust model halves for patient 1 when adding dwell positions, whereas it doubles for patient 2. These discrepancies may be caused by how much "luck" we had with the branch-and-bound tree (which also holds for the nominal optimization model). The iterative nature of our approach introduces an additional factor of "luck" in the search for an optimal solution: for one case, the procedure chose the most important scenario in an early stage, whereas more iterations were needed in another case.

For patients 1, 2 and 3 images were acquired using transrectal ultrasound imaging, whereas CT images were acquired for patients 4, 5 and 6. This results in a different delineation of the rectum: with transrectal ultrasound only the anterior wall is delineated, while the whole rectum is delineated on CT images. As a result, the rectum volumes for patients 4, 5 and 6 are approximately twice as large as for patients 1, 2 and 3. This implies that the constraint on rectum dose, $D_{10\%}(\text{Rectum}) \leq 7.2$ Gy, is less restrictive for the CT patients than for the ultrasound patients. This is inherent to using a restriction on a structure's relative volume. For our study however, it has an advantage: the results show that including rectum delineation uncertainty yields a reduction in target coverage when $D_{10\%}(\text{Rectum})$ was above 7.2 Gy for a large subset of the scenarios, whereas target coverage is not affected when $D_{10\%}(\text{Rectum}) \leq 7.2$ Gy was satisfied for most of the scenarios.

Our approach may be applicable for other body sites as well, and would require reliable data on target volume and OAR delineation uncertainties of the particular organ(s) in question. Furthermore, one may apply our method to setup uncertainties by viewing these variations as a rigid shift of the organs and hence the delineations, leaving the position of the calculation points and hence the dose rate fixed. This would allow us to combine all uncertainties into one composite uncertainty in the index set. A thorough investigation is mandatory to assess the feasibility and value of this approach. Moreover, this approach can be used not only for HDR Brachytherapy, but also for external beam radiotherapy.

The core idea of this paper on how to deal with uncertainties in index sets could also be applicable to other Operations Research problems. A first type of models are optimization models that have uncertainty in the time index. E.g., supply chain models often have uncertainty in the lead time. In the food supply chain model for the UN World Food Program (Peters et al. 2016), for example, there is much uncertainty in the lead time because of possible congestion in the harbors. A second type of models are optimization models that have different sets of restrictions for different types of customers, regions, parts, etc. An example could be inventory models with different categories of service levels, depending on the characteristics of the customer, and there is uncertainty about the required service level category for a subset of the customers in the future. Another specific example is optimizing flood protection measures for the coming centuries such that all safety levels are satisfied (Postek et al. 2016). However, for the future for a certain dike ring area there is much uncertainty about whether and when there will be a change in safety level category as a consequence of political decision making.

5. Conclusion

The worst-case robust treatment plan optimization model presented in this work for prostate HDR BT is well capable of accounting for target volume and rectum delineation uncertainties. Uncertainties in index sets can be accounted for by using a scenario-based approach. Although the treatment plan optimization model becomes too large to be solved

within a clinically acceptable amount of time, our heuristic approach reduces the calculation times to acceptable proportions for both the nominal and the robust optimization models.

Acknowledgments
We thank Ulrich Wimmert† from SonoTECH GmbH (Neu-Ulm, Germany) for providing a research version of the HDRplus software that has the ability to export the dose rate kernel matrix and the coordinates of surface points, dose calculation points and dwell positions.

Appendix. Full nominal treatment plan optimization model

A. Full nominal plan optimization model
The full treatment plan optimization model described by Gorissen et al. (2013) is the following:

$$
\begin{aligned}
(LDV) \quad \max \quad & \sum_{i \in I_P} v_i & & \\
\text{s.t.} \quad & \dot{d}_i^T t \geq v_i L & \forall i \in I_P & (9)\\
& \dot{d}_i^T t \leq L_U + (U_U - L_U)(1 - z_i) & \forall i \in I_U & (10)\\
& \sum_{i \in I_U} z_i \geq \tau_U |I_U| & & (11)\\
& t_j \leq T_{max} b_k & \forall j \in J_k, \forall k \in K & (12)\\
& b_{k_1} \leq 1 - b_{k_2} & \forall k_2 \in \Gamma(k_1), \forall k_1 \in K & (13)\\
& \sum_{k \in K} b_k \leq 20 & & (14)\\
& t \geq 0 & & \\
& v_i \in \{0,1\} & \forall i \in I_P & \\
& u_i \in \{0,1\} & \forall i \in I_\diamond, \forall \diamond \in \{R,U\} & \\
& b_k \in \{0,1\} & \forall k \in K. &
\end{aligned}
$$

The objective and constraint (9) are discussed in Section 2.2. Constraint (10) ensures that a calculation point in the urethra does not receive a dose above U_U. At most a fraction τ_U of the calculation points in the urethra may receive a dose above L_U, as is enforced by constraints (10) and (11). The variable b_k is binary, and is equal to 1 when catheter k is used. The set K is the set of possible catheter locations, and J_k is the set of dwell positions in catheter k. At most 20 catheters can be used (constraint (14)). Dwell times within catheter k can only be positive if the catheter position is used, and can never exceed a predetermined maximum dwell time T_{max} (constraint (12)). Two neighboring catheters cannot be used both, as is ensured by constraint (13), where $\Gamma(k)$ is the set of catheter locations neighboring location k. The parameter values that were used in our tests can be found in Table 10. The maximum allowed dwell time is set to 5 seconds.

Table 10 Model parameters used in the numerical experiments.

	Rectum	Urethra
$L_\diamond$	7.2 Gy	8 Gy
$U_\diamond$	10 Gy	10.6 Gy
$\tau_\diamond$	0.9	0.9

B. Algorithm procedures

Procedure "Input and initialize"

Choose values for input parameters:
Optimality gaps g_1 and g_2, $g_1 > g_2$
Fractions of calculation points $\eta_{100\%}$ and $\eta_{200\%}$
Set optimality gap to g_1.
Initialize $S_{100\%} = \{\text{nom}\}$, $S_{200\%} = \emptyset$
Solve (RC), denote solution by t^*.
$\tilde{I}_C^n := \{i \in \tilde{I}_C : \dot{d}^T t^* \geq L\}$
$\tilde{I}_C^f := \{i \in \tilde{I}_C : \dot{d}^T t^* \leq 2L\}$
if $\arg\min_{s\in S}\{V^s_{100\%}(t^*)\} \neq \text{nom}$ **then** $S_{100\%} := S_{100\%} \cup \arg\min_{s\in S}\{V^s_{100\%}(t^*)\}$
end if
if $\max_{s\in S}\{V^s_{200\%}(t^*)\} > 0.2$ **then** $S_{200\%} := \arg\max_{s\in S}\{V^s_{200\%}(t^*)\}$
end if

Procedure "Update $\tilde{I}_C^n$ and $\tilde{I}_C^f$"
Note: $\dagger = 1$ if catheters a variable, 2 if catheters are fixed.

Calculate the worst case objective value V_{true} and $V_{200\%}$ for the current solution using the complete set $\tilde{I}_C$, based on scenario sets $S_{100\%}$ and $S_{200\%}$ respectively.
if $V_{\text{true}} - V^* > g_\dagger$ **then** Remove the $\eta_{100\%}|\tilde{I}_{CTV}|$ calculation points with the lowest dose from $\tilde{I}_C^n$.
end if
if $V_{200\%} > 0.2$ **then** Remove the $\eta_{200\%}|\tilde{I}_{CTV}|$ calculation points with the highest dose from $\tilde{I}_C^f$.
end if

Procedure "Update $S_{100\%}$ and $S_{200\%}$"

if $\bar{s} := \arg\min_{s\in S}\{V_{100\%}(t^*)\} \notin S_{100\%}$ **then** $S_{100\%} := S_{100\%} \cup \bar{s}$.
end if
if $\max_{s\in S}\{V^s_{200\%}\} > 0.2$ **then** $S_{200\%} := S_{200\%} \cup \arg\max_{s\in S}\{V^s_{200\%}\}$.
end if

C. Results

C.1. DVH distributions for patients 2 up to 6

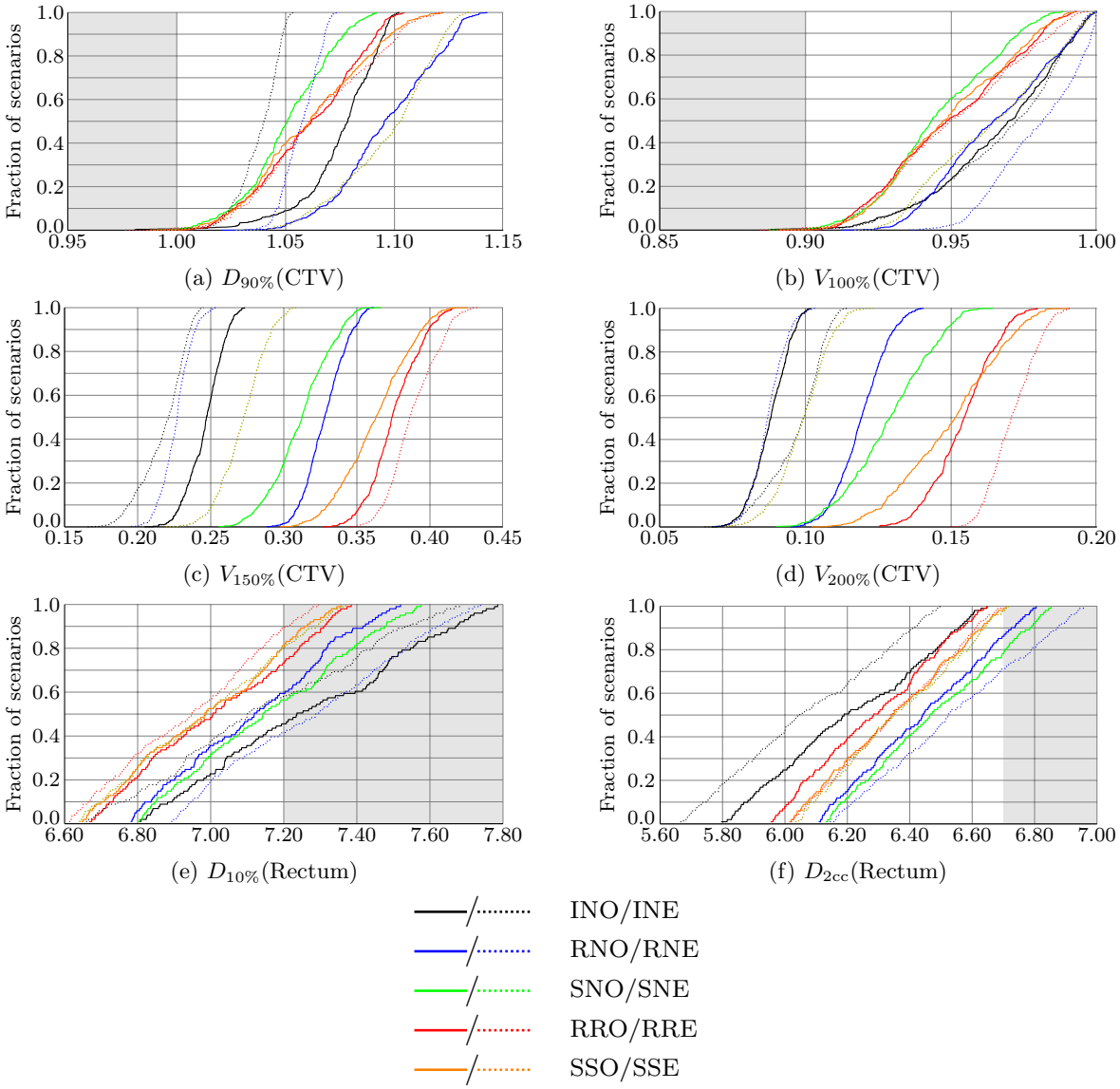

Figure 8 **DVH metrics for patient 2 obtained with various models, evaluated in the evaluation dataset. The curves for SNE and SSE coincide, as the treatment plans are identical. The gray area indicates scenarios where the DVH requirement from Table 1 was not met.**

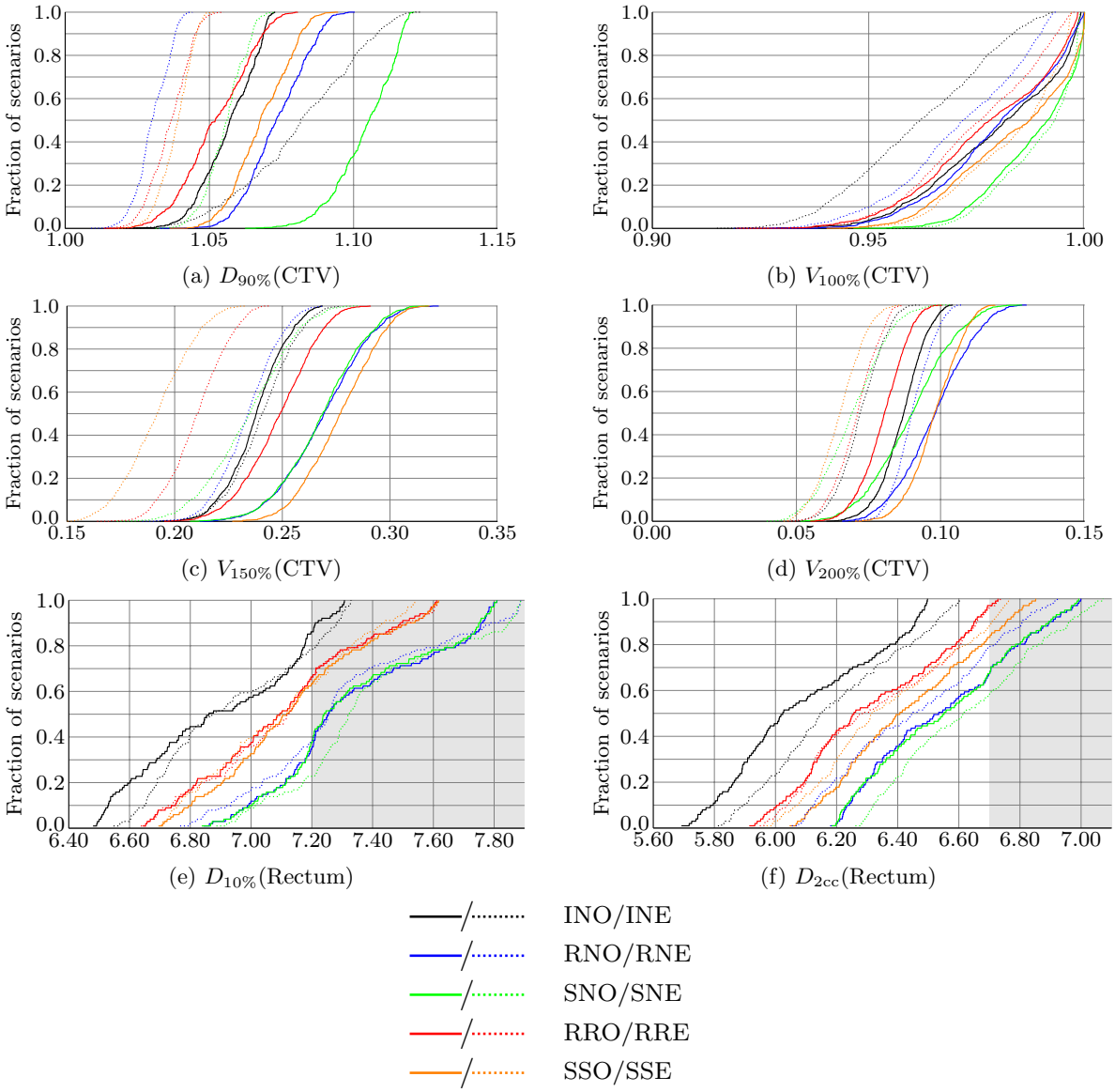

Figure 9 **DVH metrics for patient 3 obtained with various models, evaluated in the evaluation dataset. The gray area indicates scenarios where the DVH requirement from Table 1 was not met.**

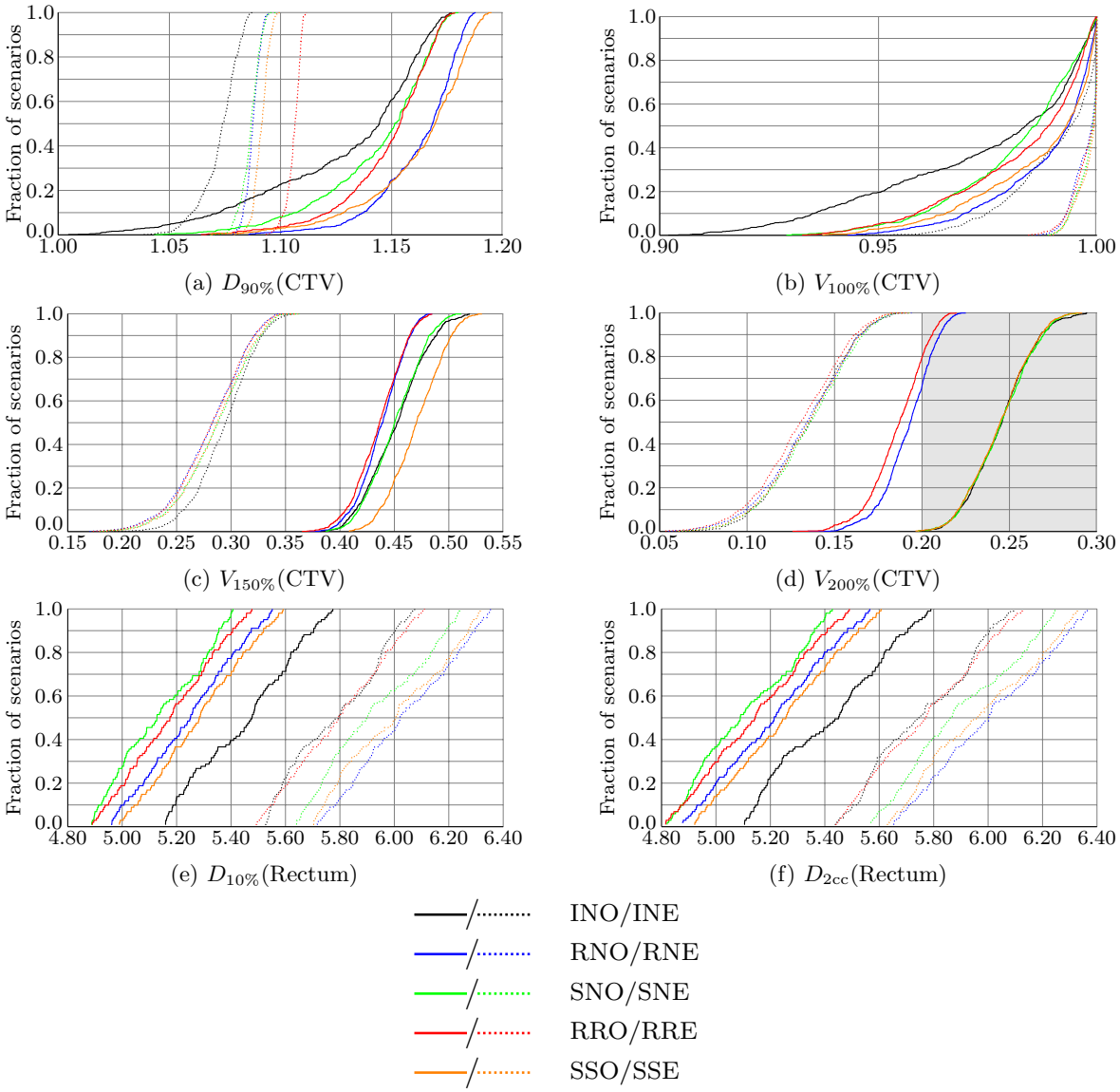

Figure 10 **DVH metrics for patient 4 obtained with various models, evaluated in the evaluation dataset. The gray area indicates scenarios where the DVH requirement from Table 1 was not met.**

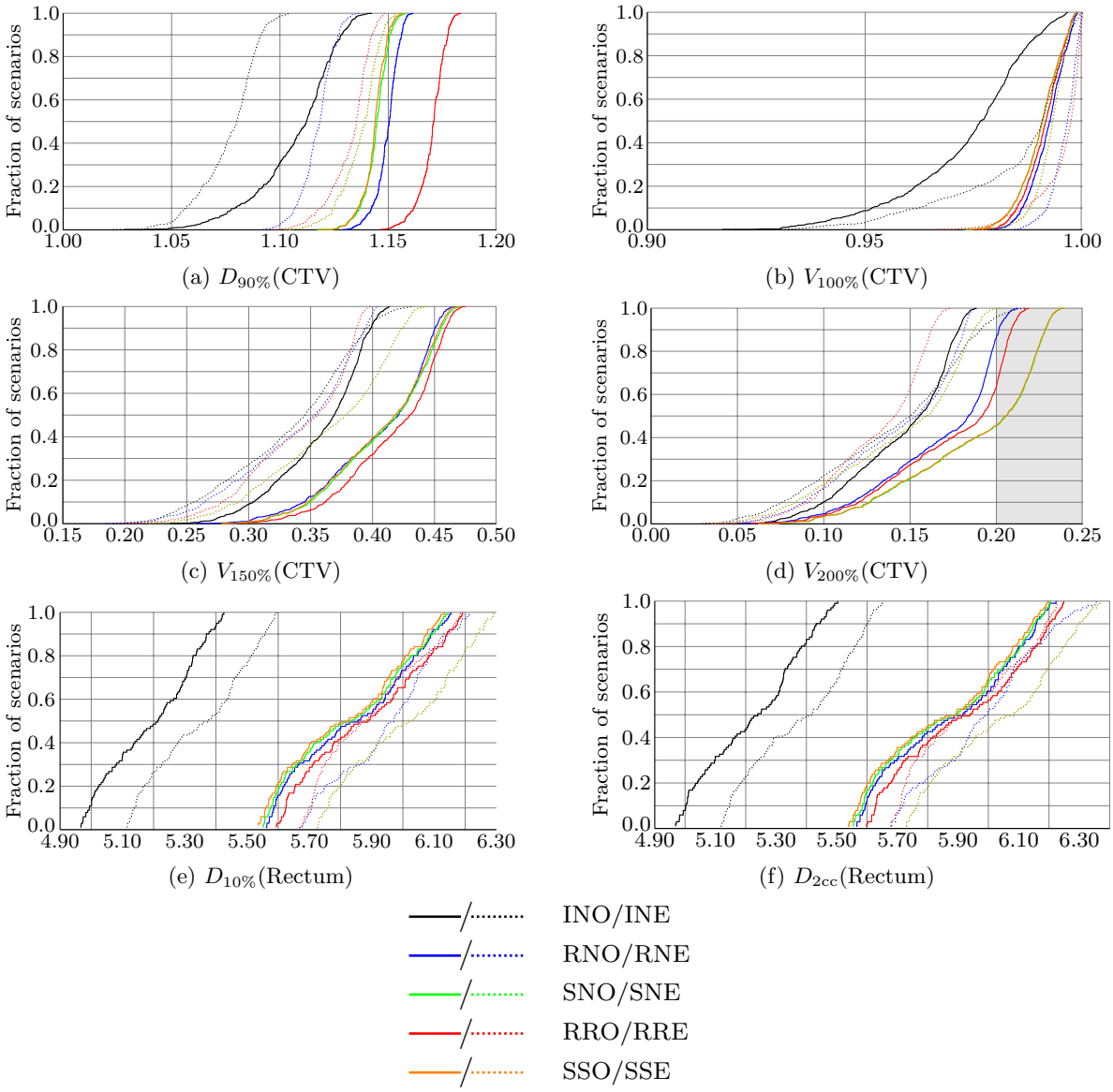

Figure 11 **DVH metrics for patient 5 obtained with various models, evaluated in the evaluation dataset. The gray area indicates scenarios where the DVH requirement from Table 1 was not met.**

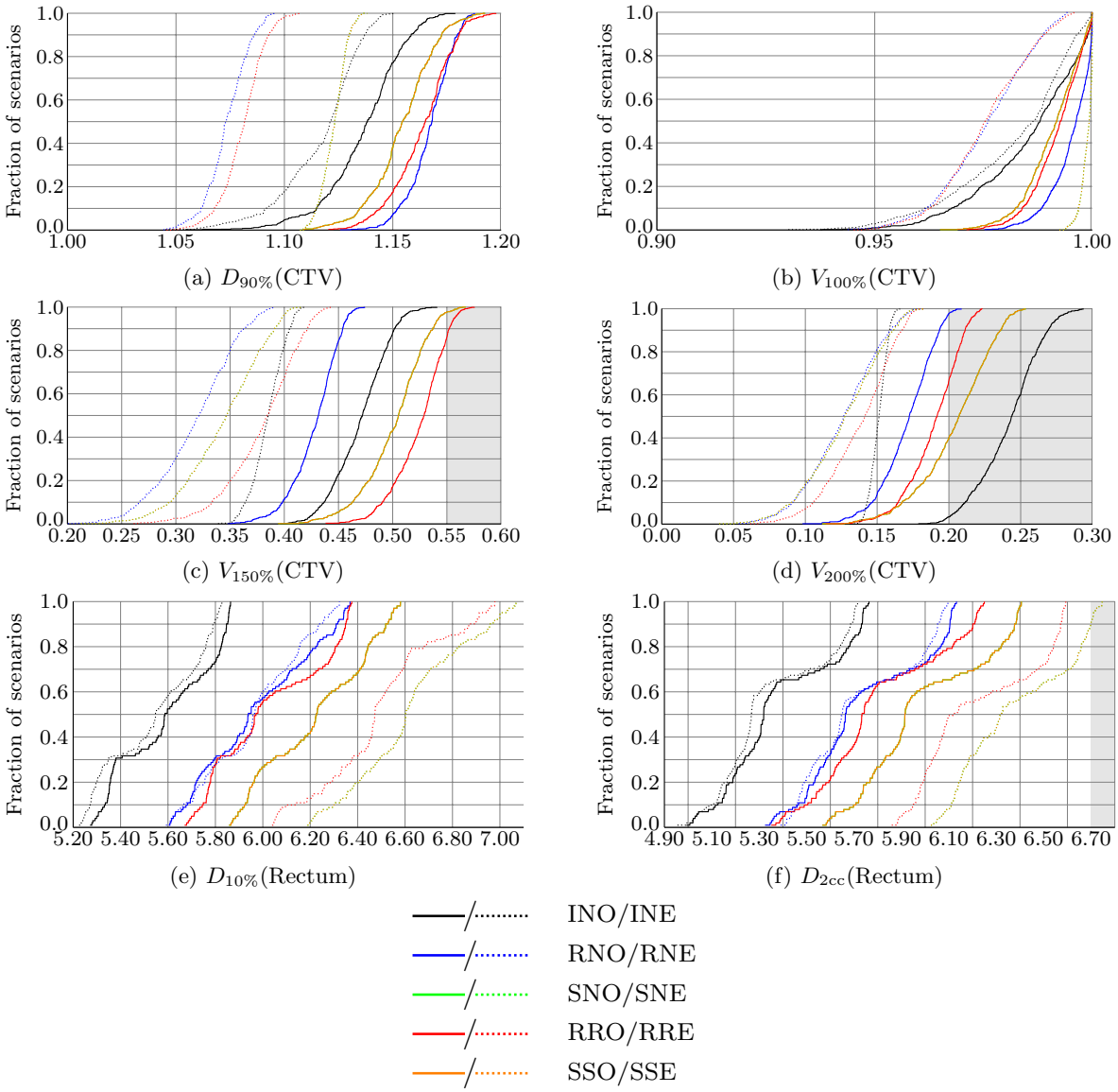

Figure 12 **DVH metrics for patient 6 obtained with various models, evaluated in the evaluation dataset. The gray area indicates scenarios where the DVH requirement from Table 1 was not met.**

C.2. DVH parameters for the urethra

Table 11 DVH parameters of the urethra for the ten optimization models, measured in Gy.

Model	DVH parameter and constraint	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6
INO	$D_{10\%}$(Urethra)≤10	9.9	9.8	9.5	9.8	9.8	9.7
	$D_{0.1cc\%}$(Urethra)≤10	10.4	10.8	10.3	10.1	10.2	9.8
RNO	$D_{10\%}$(Urethra)≤10	10.0	9.9	9.9	10.0	10.0	10.0
	$D_{0.1cc\%}$(Urethra)≤10	10.5	10.4	10.6	10.3	10.6	10.2
RRO	$D_{10\%}$(Urethra)≤10	10.0	10.0	9.8	10.0	10.0	9.9
	$D_{0.1cc\%}$(Urethra)≤10	10.4	10.5	10.6	10.5	10.5	10.1
SNO	$D_{10\%}$(Urethra)≤10	9.9	10.0	9.9	10.0	10.0	10.0
	$D_{0.1cc\%}$(Urethra)≤10	10.3	10.5	10.5	10.2	10.4	10.3
SSO	$D_{10\%}$(Urethra)≤10	9.9	10.0	10.0	10.0	10.0	10.0
	$D_{0.1cc\%}$(Urethra)≤10	10.3	10.5	10.6	10.2	10.4	10.3
INO	$D_{10\%}$(Urethra)≤10	9.9	9.8	9.8	9.8	9.6	9.6
	$D_{0.1cc\%}$(Urethra)≤10	10.4	10.6	10.5	9.9	10.0	9.7
RNO	$D_{10\%}$(Urethra)≤10	9.9	9.8	9.4	10.0	10.0	9.9
	$D_{0.1cc\%}$(Urethra)≤10	10.6	10.3	10.2	10.2	10.4	10.1
RRO	$D_{10\%}$(Urethra)≤10	10.0	9.8	9.6	10.0	10.0	9.9
	$D_{0.1cc\%}$(Urethra)≤10	10.5	10.0	10.5	10.2	10.4	10.2
SNO	$D_{10\%}$(Urethra)≤10	10.0	9.9	9.8	10.0	10.0	10.0
	$D_{0.1cc\%}$(Urethra)≤10	10.3	10.5	10.6	10.2	10.5	10.3
SSO	$D_{10\%}$(Urethra)≤10	10.0	9.9	9.7	9.9	10.0	10.0
	$D_{0.1cc\%}$(Urethra)≤10	10.3	10.5	10.6	10.0	10.5	10.3

D. Changing the algorithm's parameters - detailed results

Table 12 The effect of the allowed optimality gap g_2 on solution time, objective function value and worst case $V_{100\%}$(CTV), both for the optimization scenarios and for the simulations.

	Gap (%)	Solution time	Objective value	Worst case V_{100}%(CTV) over scenarios	Worst case V_{100}%(CTV) over simulations
Patient 1	0.3	2389	98.9	88.1	86.4
	0.4	571	98.9	87.2	85.8
	0.5	444	98.6	85.2	84.9
	0.75	310	98.6	86.0	84.8
Patient 3	0.025	176	99.8	96.1	92.7
	0.05	175	99.8	96.1	92.7
	0.1	165	99.7	95.8	92.2
	0.2	138	99.7	95.6	92.1
	0.3	112	99.7	95.6	92.1
	0.4	74	99.4	90.0	87.8

Table 13 The effect of different values for $\eta_{100\%}$ and $\eta_{200\%}$ on the solution times,objective function value and the worst case $V_{100\%}$(CTV), both for the optimization scenarios and for the simulations. The results shown are for patient 1.

$\eta_{100\%}$	$\eta_{200\%}$	Solution time	Objective value	Worst case V_{100}%(PTV) over scenarios	Worst case V_{100}%(PTV) over simulations
0.05	0.05	1274	98.8	87.9	85.7
0.05	0.10	915	98.9	87.3	86.0
0.05	0.15	904	98.9	87.3	86.0
0.10	0.05	481	98.7	86.4	85.0
0.10	0.10	571	98.9	87.2	85.8
0.10	0.15	1811	98.9	87.3	85.9
0.15	0.05	472	98.7	86.4	85.0
0.15	0.10	1819	98.9	87.3	85.9
0.15	0.15	1798	98.9	87.3	85.9

References

Balvert, M, B Gorissen, D den Hertog, A Hoffmann. 2015. Dwell time modulation restrictions do not necessarily improve treatment plan quality for prostate HDR brachytherapy. *Physics in Medicine and Biology* **60** 537–548.

Ben-Tal, A, L El Ghaoui, A Nemirovski. 2009. *Robust Optimization*. Princeton Series in Applied Mathematics, Princeton University Press.

Bienstock, D, N Özbay. 2008. Computing robust basestock levels. *Discrete Optimization* **5** 389–414.

Bohoslavsky, R, M Witte, T Janssen, M van Herk. 2013. Probabilistic objective functions for margin-less IMRT planning. *Physics in Medicine and Biology* **58** 3563–3580.

Bortfeld, T, T Chan, A Trofimov, J Tsitsiklis. 2008. Robust management of motion uncertainty in intensity-modulated radiation therapy. *Operations Research* **56** 1461–1473.

Chan, T, T Bortfeld, J Tsitsiklis. 2006. A robust approach to IMRT optimization. *Physics in Medicine and Biology* **51** 2567–2583.

Chen, W, J Unkelbach, A Trofimov, T Madden, H Kooy, T Bortfeld, D Craft. 2011. Including robustness in multi-criteria opimization for intensity modulated proton therapy. *Physics in Medicine and Biology* **57** 591–608.

Chu, M, Y Zinchenko, S Henderson, M Sharpe. 2005. Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty. *Physics in Medicine and Biology* **50** 5463–5477.

De Boeck, L., J. Beliën, W. Egyed. 2014. Dose optimization in high-dose-rate brachytherapy: a literature review of quantitative models from 1990-2010. *Operational Research for Health Care* **3** 80–90.

De Brabandere, M., P. Hoskin, K. Haustermans, F. Van den Heuvel, F.-A. Siebert. 2012. Prostate post-implant dosimetry: interobserver variability in seed localisation, contouring and fusion. *Radiotherapy and Oncology* **104** 192–198.

Deist, T. M., B. L. Gorissen. 2016. High-dose-rate prostate brachytherapy inverse planning on dose-volume criteria by simulated annealing. *Physics in Medicine and Biology* **61** 1155–1170.

Fredriksson, A. 2012. A characterization of robust radiation therapy treatment planning methods - from expected value to worst case optimization. *Medical Physics* **39** 5169–5181.

Fredriksson, A. 2013. Robust optimization of radiation therapy accounting for geometric uncertainty. Ph.D. thesis, KTH Engineering Sciences, Stockholm, Sweden.

Fredriksson, A, R Bokrantz. 2014. A critical evaluation of worst case optimization methods for robust intensity-modulated proton therapy planning. *Medical Physics* **41** 081701.

Fredriksson, A, A Forsgren, B Hårdemark. 2011. Minimax optimization for handling range and setup uncertainties in proton therapy. *Medical Physics* **38** 1672.

Gorissen, B, D den Hertog, A Hoffmann. 2013. Mixed integer programming improves comprehensibility and plan quality in inverse optimization of prostate HDR brachytherapy. *Physics in Medicine and Biology* **58** 1041–1057.

Granero, D, J Pérez-Calatayud, E Casal, F Ballester, J Venselaar. 2006. A dosimetric study on the Ir-192 high dose rate Flexisource. *Medical Physics* **33** 4578–4582.

Holm, Å, T Larsson, Å Tedgren. 2013. A linear programming model for optimizing HDR brachytherapy dose distributions with respect to the mean dose in the DVH-tail. *Medical Physics* **40** 081795.

Hoskin, P, K Motohashi, P Bownes, L Bryant, P Ostler. 2007. High dose rate brachytherapy in combination with external beam radiotherapy in the radical treatment of prostate cancer: initial results of a randomised phase three trial. *Radiotherapy and Oncology* **84** 114–120.

International Commission on Radiation Units and Measurements. 1999. *ICRU 62. Prescribing, recording and reporting photon beam therapy*.

Kirisits, C, M Rivard, D Baltas, F Ballester, M De Brabandere, R Van der Laarse, Y Niatsetski, P Papagiannis, T Hellebust, J Perez-Calatayud, K Tanderup, J Venselaar, F Siebert. 2014. Review of clinical brachytherapy uncertainties: Analysis guidelines of GEC-ESTRO and the AAPM. *Radiotherapy and Oncology* **110** 199–212.

Lessard, E, J Pouliot. 2001. Inverse planning anatomy-based dose optimization for HDR-brachytherapy of the prostate using fast simulated annealing algorithm and dedicated objective function. *Medical Physics* **28** 773–779.

Liu, W, X Zhang, Y Li, R Mohan. 2012. Robust optimization of intensity modulated proton therapy. *Medical Physics* **39** 1079–1091.

Nath, R., L. Anderson, G. Luxton, K. Weaver, J. Williamson, A. Meigooni. 1995. Dosimetry of interstitial brachytherapy sources: recommendations of the AAPM radiation therapy committee task group no. 43. *Medical Physics* **22** 209–234.

Olafsson, A, S Wright. 2006. Efficient schemes for robust IMRT treatment planning. *Physics in Medicine and Biology* **51** 5621–5642.

Peters, K, H Fleuren, D den Hertog, M Kavelj, S Silva, R Goncalves, O Ergun, M Soldner. 2016. The nutritious supply chain: Optimizing humanitarian food aid. *CentER Discussion Paper* .

Postek, K, D den Hertog, J Kind, C Pustjens. 2016. Adjustable robust strategies for flood protection. *CentER Discussion Paper* .

Rylander, S, S Buus, E M Pedersen, L Bentzen, K Tanderup. 2017. Dosimetric impact of contouring and needle reconstruction uncertainties in US-, CT- and MRI-based high-dose-rate prostate brachytherapy treatment planning. *Radiothera* **123** 125–132.

Siauw, T, A Cunha, A Atamtürk, I Hsu, J Pouliot, K Goldberg. 2011. IPIP: A new approach to inverse planning for HDR brachytherapy by directly optimizing dosimetric indices. *Medical Physics* **38** 4045–4051.

Smith, W, C Lewis, G Bauman, G Rodrigues, D D'Souza, R Ash, D Ho, V Venkatesan, D Downey, A Fenster. 2007. Prostate volume contouring: a 3D analysis of segmentation using 3DTRUS, CT and MR. *International Journal of Radiation Oncology* **67** 1238–1247.

Tanderup, K, R Pötter, J Lindegaard, D Berger, A Wambersie, C Kirisits. 2010. PTV margins should not be used to compensate for uncertainties in 3D image guided intracavitary brachytherapy. *Radiotherapy and Oncology* **97** 495–500.

Unkelbach, J, U Ulfke. 2004. Inclusion of organ movements in IMRT treatment planning via inverse planning based on probability distributions. *Physics in Medicine and Biology* **49** 4005–4029.

Van Herk, M. 2004. Errors and margins in radiotherapy. *Seminars in Radiation Oncology* **14** 52–64.

Van Herk, M., P. Remeijer, J. V. Lebesque. 2002. Inclusion of geometric uncertainties in treatment plan evaluation. *International Journal of Radiation Oncology Biology Physics* **52** 1407–1422.

Villeirs, G. M., K. Van Vaerenbergh, L. Vakaet, S. Bral, F. Claus, W. J. De Neve, K. L. Verstraete, G. O. De Meerleer. 2005. Interobserver delineation variation using CT versus combined CT + MRI in intensity-modulated radiotherapy for prostate cancer. *Strahlentherapie und Onkologie* **181** 424–430.

Weiss, E, C Hess. 2003. The impact of gross tumor volume (GTV) and clinical target volume (CTV) definition on the total accuracy in radiotherapy theoretical aspects and practical experiences. *Strahlentherapie und Onkologie* **179** 21–30.

Yamada, Y, L Rogers, D Demanes, G Morton, B Prestidge, J Pouliot, G Cohen, M Zaider, M Ghilezan, I Hsu. 2012. American brachytherapy society consensus guidelines for high-dose-rate prostate brachytherapy. *Brachytherapy* **11** 20–32.