ELSEVIER

BRACHYTHERAPY

Brachytherapy 18 (2019) 396–403

Physics

Evaluation of bi-objective treatment planning for high-dose-rate prostate brachytherapy—A retrospective observer study

Stefanus C. Maree[1,*], Ngoc Hoang Luong[2], Ernst S. Kooreman[1], Niek van Wieringen[1], Arjan Bel[1], Karel A. Hinnen[1], Henrike Westerveld[1], Bradley R. Pieters[1], Peter A.N. Bosman[2,3], Tanja Alderliesten[1]

[1]*Department of Radiation Oncology, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands*
[2]*Life Sciences and Health Research Group, Centrum Wiskunde & Informatica, Amsterdam, The Netherlands*
[3]*Algorithmics group, Department of Software Technology, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands*

ABSTRACT

PURPOSE: Bi-objective treatment planning for high-dose-rate prostate brachytherapy is a novel treatment planning method with two separate objectives that represent target coverage and organ-at-risk sparing. In this study, we investigated the feasibility and plan quality of this method by means of a retrospective observer study.

METHODS AND MATERIALS: Current planning sessions were recorded to configure a bi-objective optimization model and to assess its applicability to our clinical practice. Optimization software, GOMEA, was then used to automatically generate a large set of plans with different trade-offs in the two objectives for each of 18 patients treated with high-dose-rate prostate brachytherapy. From this set, five plans per patient were selected for comparison to the clinical plan in terms of satisfaction of planning criteria and in a retrospective observer study. Three brachytherapists were asked to evaluate the blinded plans and select the preferred one.

RESULTS: Recordings demonstrated applicability of the bi-objective optimization model to our clinical practice. For 14/18 patients, GOMEA plans satisfied all planning criteria, compared with 4/18 clinical plans. In the observer study, in 53/54 cases, a GOMEA plan was preferred over the clinical plan. When asked for consensus among observers, this ratio was 17/18 patients. Observers highly appreciated the insight gained from comparing multiple plans with different trade-offs simultaneously.

CONCLUSIONS: The bi-objective optimization model adapted well to our clinical practice. GOMEA plans were considered equal or superior to the clinical plans. In addition, presenting multiple high-quality plans provided novel insight into patient-specific trade-offs.

Keywords: Brachytherapy; HDR; Prostate neoplasm; Treatment planning; Bi-objective optimization; Observer study

Introduction

Current commercial treatment planning methods for high-dose-rate (HDR) prostate brachytherapy (BT) such as inverse planning simulated annealing (IPSA) (1, 2), hybrid inverse treatment planning and optimization (HIPO) (3), or graphical optimization (4) present a single plan to the planner. This plan can then iteratively be adapted by changing dwell times, drag-and-drop of isodose lines, or by changing the parameters of the planning method. After each modification, the planner assesses the plan quality by a set of planning criteria on

Received 8 August 2018; received in revised form 8 November 2018; accepted 28 December 2018.

Conflict of interest: T.A., A.B., B.R.P., and P.A.N.B. are involved in projects supported by Elekta. Elekta had no involvement in the study design, data collection, analysis and interpretation, and writing of the article.

* Corresponding author. Department of Radiation Oncology, Amsterdam UMC, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. Tel.: +31 20 56 64446.

E-mail address: s.c.maree@amc.uva.nl (S.C. Maree).

1538-4721/$ - see front matter
https://doi.org/10.1016/j.brachy.2018.12.010

the dose-volume indices (DVIs) stated in the clinical protocol and by visual inspection of the dose distribution. Obtaining acceptable plans by this procedure requires experience and is time-consuming, often taking more than 30 min (4–6).

To efficiently obtain high-quality plans using inverse planning methods, the underlying optimization model should closely match the planning criteria, while computation time is still clinically acceptable. The optimization model on which IPSA and HIPO are based—a dose-penalty model—is fast to compute but does not always result in plans that adhere to the planning criteria, even if such plans do exist (7, 8). Other methods (5, 9, 10) do model the treatment planning problem based on the DVIs of the planning criteria. However, similar to IPSA and HIPO, multiple criteria are then combined into a single model by a weighting that the planner needs to set. Tuning these weights is a patient-dependent and nontrivial task, which makes treatment planning a difficult and time-consuming trial-and-error process (11, 12).

To overcome tuning of weights, the treatment planning problem can be formulated as a multi-objective optimization problem, where each planning criteria is an objective (13). Typically, there are five or more objectives, depending on the treatment site and clinical protocol (14). The optimum of a multi-objective optimization model is not a single treatment plan, but a large set of plans, all with a different trade-off between the objectives. A single preferred plan then needs to be selected from this set, either manually or automatically. However, considering all planning criteria as separate objectives results in many objectives, which makes fast computing of all best trade-off plans computationally infeasible (15). In practice, an interactive method can be used by which optimization is steered by the planner to get to a desirable plan (16–18). Although more intuitive than setting weights, it is hard to get intuition about the nature of underlying trade-offs this way.

Alternatively, approaches exist to navigate the set of trade-off solutions without first computing them all, but approximations of the planning criteria (i.e., dose-penalty models rather than DVI models) and plan interpolations are then required for efficiency reasons, making it more difficult to obtain plans that are in line with the planning criteria (7, 8, 19, 20).

The novel treatment planning method that is used in the present study models the treatment planning problem as a bi-objective optimization problem, with only two objectives. One objective is based on target coverage and the other on organ-at-risk (OAR) sparing (21). The objectives are directly based on the DVIs stated in the clinical protocol. As there are only two objectives, first computing a set of high-quality trade-off plans is computationally tractable, and visualization of these plans as a trade-off curve is straightforward. This reduces treatment planning to a decision-making process of selecting the preferred plan from this trade-off curve.

In this work, we evaluate the use and plan quality of bi-objective treatment planning for our clinical practice in a retrospective observer study.

Methods and materials

Patient and treatment characteristics

Between February 2015 and April 2017, 18 prostate cancer patients were treated in our clinic according to the protocol in Table 1, receiving single-dose HDR BT of 13 Gy a week after external beam radiation treatment with a dose schedule of 20 × 2.2 Gy. Median age at time of treatment was 68 (range: 58–84) years; Gleason score was between 6 and 9 (ISUP grade group 1–5). The median urinary flow rate was 16.3 (range: 8.5–34.8) mL/s, and the median prostate volume defined by MRI after catheter placement was 31.9 (range: 21.1–69.3) cm^3. A median of 16 (range: 14–20) catheters were implanted with a source step of 2.5 mm, totaling to a median of 413 (range: 250–668) dwell positions. Catheter implantation was performed using transrectal ultrasound under general or epidural anesthesia according to a preplan, made in the operation theater based on ultrasound imaging, in Oncentra Prostate (Elekta AB, Stockholm, Sweden) (22). Visibility of the urethra was enhanced by a transurethral catheter with a bladder balloon.

After implantation, three orthogonal pelvic T2-weighted turbo spin echo MRIs (Ingenia 3T Philips Healthcare, Best, The Netherlands) with an in-plane resolution of 0.6 × 0.7 mm and 3.0 mm slice thickness with 0.3 mm gap were acquired and used for treatment planning. Imaging was taken with the patients lying on their back and legs flat,

Table 1
High-dose-rate brachytherapy protocol of a single planning-aim dose of 13 Gy

Volume	Use	Coverage criteria		Sparing criteria	
Prostate	Target	$V_{100\%} > 95\%$ (42 Gy $EQD2_3$)	$D_{90\%} > 100\%$ (42 Gy $EQD2_3$)	$V_{150\%} < 50\%$ (89 Gy $EQD2_3$)	$V_{200\%} < 20\%$ (150 Gy $EQD2_3$)
Vesicles	Target	$V_{80\%} > 95\%$ (28 Gy $EQD2_3$)			
Bladder	OAR			$D_{1cm^3} < 86\%$ (32 Gy $EQD2_3$)	$D_{2cm^3} < 74\%$ (24 Gy $EQD2_3$)
Rectum	OAR			$D_{1cm^3} < 78\%$ (27 Gy $EQD2_3$)	$D_{2cm^3} < 74\%$ (24 Gy $EQD2_3$)
Urethra	OAR			$D_{0.1cm^3} < 110\%$ (50 Gy $EQD2_3$)	

OAR = organ at risk; V = Volume indices in percent of the planning-aim dose; D = Dose indices in volume percentage or absolute volume (cm^3) and units gray (Gy).

similar to the treatment position. These images were loaded into Oncentra Brachy (version 4.3–4.5, Elekta AB, Stockholm, Sweden) and used for catheter reconstruction and delineation of the regions of interest.

Initial plans of patients treated before mid-2015 were created using IPSA, initial plans after mid-2015 were created with HIPO. Both IPSA and HIPO were run with a standard parameter set, that is, a class solution, as provided in the Supplementary Material Figs. 7 and 8. Plans were then manually fine-tuned using graphical optimization. Next, quality assurance checks were done by a medical physicist. Finally, the plans were assessed for clinical acceptability by a physician using the criteria in Table 1 and by visual inspection of the dose distribution.

Analysis of clinical planning sessions

To gain insight into the current planning process in our clinic and into the applicability of the bi-objective optimization model (21), clinical planning sessions were filmed, and changes in DVIs over time were recorded during the manual graphical optimization of five of the 18 patients.

It was measured how many of the changes were dedicated to improving the DVIs and how many focused on other aspects not explicitly mentioned in the clinical protocol. Based on this analysis, the bi-objective optimization model was configured (23, 24).

Configuring the bi-objective optimization model

The bi-objective optimization model groups the planning criteria (Table 1) into one coverage objective and one sparing objective. We configured the model as follows:

$$LCI = \min\left\{V_{100\%}^{prostate} - 95, V_{80\%}^{vesicles} - 95\right\},$$

$$LSI = 13Gy \times \min\left\{\begin{array}{c} 86 - D_{1\text{cm}3}^{bladder}, 74 - D_{2\text{cm}3}^{bladder}, \\ 78 - D_{1\text{cm}3}^{rectum}, 74 - D_{2\text{cm}3}^{rectum}, \\ 110 - D_{0.1\text{cm}3}^{urethra} \end{array}\right\}.$$

The least coverage index (LCI) was constructed by combing the coverage criteria $V_{100\%}^{prostate}$ and $V_{80\%}^{vesicles}$ in a worst-case manner. An LCI = 1.2% should be read as "the worst covered target is covered 1.2% more than its aim and the other target has a higher coverage." Thus, when maximizing the LCI, a certain level of coverage for both $V_{100\%}^{prostate}$ and $V_{80\%}^{vesicles}$ is guaranteed. The criterion $D_{90\%}^{prostate} > 100\%$ has been left out of the model, as it is automatically satisfied when $V_{100\%}^{prostate} > 90\%$. Moreover, explicitly maximizing it would lead to dose escalation, and it is currently unclear whether this is desirable (25).

The least sparing index (LSI) was constructed in a similar worst-case approach from the sparing criteria (Table 1). An LSI = 1.4 Gy should be read as "the worst spared OAR is spared 1.4 Gy more than its criterion, and all other OARs are spared even more." Thus, the LSI should also be maximized, and when LSI>0 Gy holds, all sparing criteria are satisfied.

The criteria $V_{150\%}^{prostate} < 50\%$ and $V_{200\%}^{prostate} < 20\%$ are the only sparing criteria based on volume indices. Directly adding them to the LSI would result in a comparison of indices with a different unit (percentage of volume compared with percentage of prescribed dose). Analysis of clinical plans, see Supplementary Table 2, showed that the criteria $V_{150\%}^{prostate} < 50\%$ and $V_{200\%}^{prostate} < 20\%$ were never violated. As the LSI uses a worst-case approach, and as these criteria were never violated, there would be no effect of leaving them out of the objectives. To guarantee that optimization generates plans that satisfy these two criteria, all plans that violate them are automatically discarded.

Automatic bi-objective BT planning

For each patient, a large set of high-quality plans was automatically generated by optimizing plans under the bi-objective model. For this, patient DICOM RT-Struct and RT-Dose files were exported from Oncentra Brachy and processed by our in-house developed TG-43 (26, 27) dose engine. As starting point for the optimization, all dwell positions were activated within a 5-mm margin of the targets (i.e., prostate and vesicles), excluding positions within a 1-mm margin of the urethra. Next, the dwell times associated with these dwell positions were initialized with a randomly chosen value between 0 and 1 s. The aim of treatment planning is then to optimize these dwell times.

The bi-objective model is nonconvex, nonlinear, and nonsmooth and to optimize plans according to it, a state-of-the-art multi-objective evolutionary algorithm, GOMEA (23, 28), was used. In GOMEA, the fact that the dose distribution can be quickly updated when only few dwell times change is exploited. A run of GOMEA was limited to 1 h on a low-end processor. In a 1-h run, GOMEA produced a large set of 100–1000 treatment plans, all with different LCI/LSI trade-offs (23, 24). As GOMEA is a stochastic algorithm, it was run 30 times to assess its variance in final results, which was shown to be small (23). From these 30 runs, a single set of plans was constructed by only retaining plans that exhibit the best trade-offs in LCI and LSI. Further research (29, 30) showed that using a Graphics Processing Unit that is available in modern planning machines, the same results can be obtained within a matter of minutes.

For each patient, the resulting set of plans consisted of hundreds of high-quality plans. Presenting all these plans in an observer study was infeasible as software to quickly navigate through and compare that many plans is not yet available. We therefore manually selected five plans (see Fig. 1) for further analysis according to the following strategy: left and right of LCI = 0, above and below LSI = 0, and the plan in the middle. These plans are labeled from small to large LCI values as high sparing, sparing, coverage/sparing, coverage, and high coverage. These five selected GOMEA plans per patient were compared to

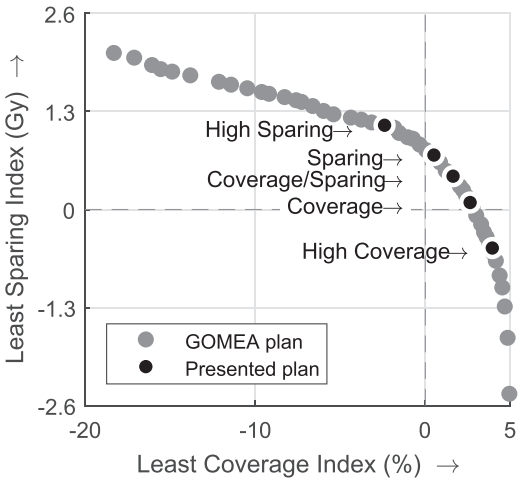

Fig. 1. Example of a trade-off curve for one patient. Each dot represents a treatment plan obtained by GOMEA. The axes measure the objectives of the bi-objective model. Both objectives should be maximized. Plans in the upper right corner, where the values for the Least Coverage Index and the Least Sparing Index are larger than zero, satisfy all planning criteria (Table 1). Large black dots illustrate the plans presented in the observer study.

clinical plans in terms of their LCI and LSI values. Differences in DVIs and LSI and LCI values were tested for significance using a Wilcoxon matched-pair signed rank test ($\alpha = 0.05$).

Observer study setup

A retrospective observer study was conducted with three physicians responsible in our clinic for HDR prostate BT. Each observer was individually presented with six plans per patient: the plan that was clinically used to treat the patient, and the five selected GOMEA plans, without identifying which plan was which. The DVIs of the six plans were presented in a single overview, similar to Oncentra Brachy. Then, the observer could inspect the corresponding dose distributions in Oncentra Brachy, one at a time.

Each observer was asked the following questions: What is the preferred plan to treat the patient with? Which plans are clinically acceptable? What plan would you dismiss?

After answering these questions, additional patient information was provided, and the observer was asked if this changed their opinion on the preferred plan. The additional information comprised patient age and Gleason score for all patients. Urinary flow rate, available biopsy information, tumor location information, and potential tumor invasion of the seminal vesicles was available for the 12 most recent patients. Furthermore, for the eight most recent patients, an additional diffusion-weighted MRI with tumor location information was added as it had been available clinically after a recent adaptation of the clinical workflow.

Additional patient information was provided to mimic clinical practice as much as possible. IPSA or HIPO do not take this additional patient information into account when initialized by a standard parameter set, but this information is known to the planner during manual optimization, and this information is thus incorporated into clinical plans. Similarly, GOMEA does not take this additional information into account during optimization, but a planner might while selecting the desired GOMEA plan. To investigate how this information changed decision-making, we chose for this setup, where the questions are repeated after the first evaluation.

It was recorded which aspects the observers assessed that were not mentioned in the clinical protocol and how they approached decision-making.

This setup of presenting multiple GOMEA plans was chosen because it provides additional insight into the patient-specific trade-off, which can be used during decision-making. The clinical plan was added to compare quality of GOMEA plans and to assess clinical acceptance. This setup does however introduce a potential bias toward selecting a GOMEA plan. We investigated statistical significance of the preferred plan by comparing against the null hypothesis that all plans are equally likely to be selected, at $\alpha = 0.05$. Because the number of observers is too low to perform an observer variability study, p-values are reported per observer.

One week after the observer study, a consensus meeting was held where patients were discussed for which each observer selected a different preferred plan. These three plans were again blindly presented, and observers were asked whether they could agree on a single preferred plan.

Results

Analysis of clinical planning sessions

The five filmed planning sessions lasted for a median of 33 min (range: 9–48), with an average of four drag-and-drop steps per minute, totaling to 525 modifications. See Supplementary Figs. 1–5 for the transcripts. Modifications were being performed on an iterative basis, focusing on two conflicting criteria: 66% of the modifications on prostate vs. urethra, 20% on vesicles vs. bladder, and 1% on prostate vs. rectum. Fifty percent of the changes were made to improve the least spared/covered volume. Two aspects, not included in the clinical protocol, that were assessed, were improving dose homogeneity and reducing hotspots, that is, volumes with dose higher than 200% of the prescribed dose, both in number and in size. For most of the manual plan optimization time, both LCI and LSI had a negative value. Over time, they were improved iteratively, which corresponds to a zigzag pattern in the bi-objective representation (Fig. 2).

Automatic bi-objective BT planning compared with clinical plans

Figure 3 shows the trade-off curves obtained with GOMEA, together with the clinical plan and the five plans that were selected from the front for comparison. For each patient, an overview of the DVIs associated with the clinical plan and the five selected plans is given in Supplementary Table 2.

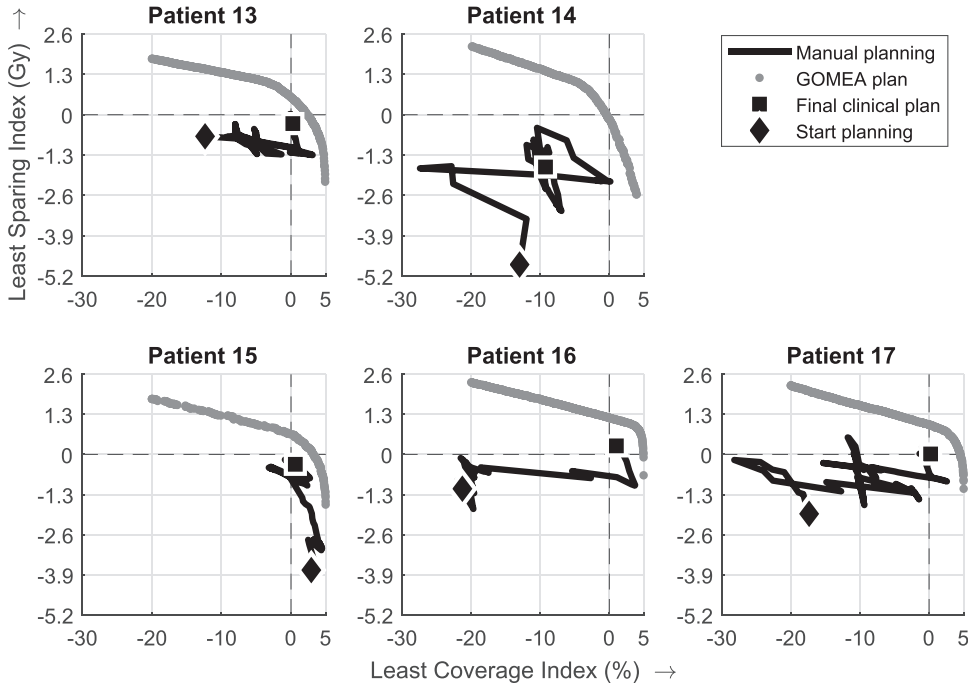

Fig. 2. Manual planning process using graphical optimization visualized in the bi-objective representation for the five recordings. The trade-off curves obtained with GOMEA are shown for comparison. Transcripts of the planning processes are given in Supplementary Figs 1–5.

For all patients, the GOMEA plans had simultaneously a better LCI and LSI than the clinical plan. For four patients (10, 11, 16, and 17), the clinical plan satisfied the clinical protocol, while GOMEA plans satisfied the clinical protocol for 14/18 of the patients. For four patients (4, 7, 14, and 18), neither the plans optimized by our method nor the clinical plan satisfied all planning criteria, caused by an unfavorable implant geometry. For some patients, the LCI value of the clinical plan is small because of a low vesicle coverage. For all clinical plans, the $V_{100\%}^{prostate}$ was ≥93.3 %.

From the set of five GOMEA plans per patient, GOMEA plans with a similar or better LSI as the clinical plan had an LCI that is 3.5% larger than the LCI of the clinical plan,

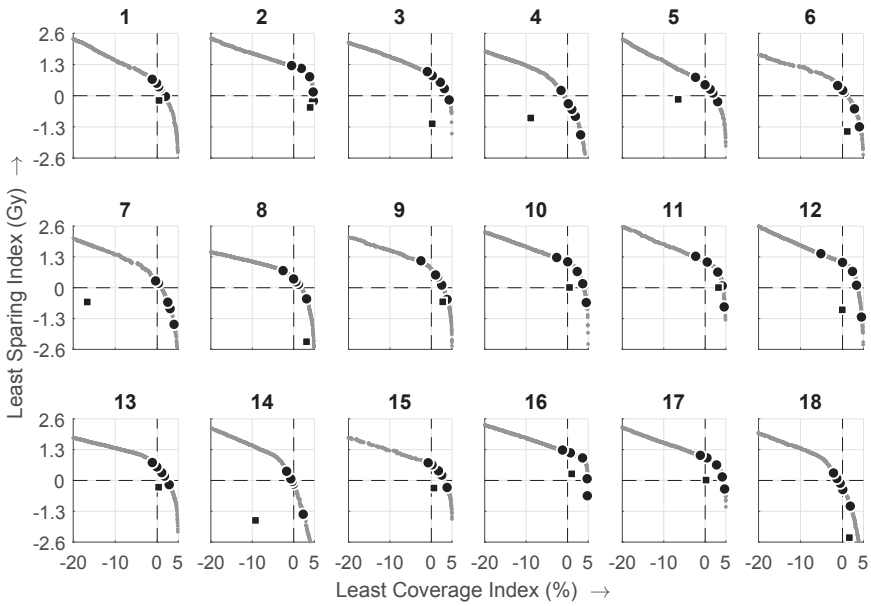

Fig. 3. Trade-off curves for all 18 patients in gray small dots. Black squares represent the clinical plans. Black circles are the GOMEA plans presented in the observer study, representing from left to right on the trade-off curve the selected high-sparing, sparing, coverage/sparing, coverage, and high-coverage plan.

averaged over all patients (range −0.6–14.7%, SD 4.3%, $p < 0.05$). GOMEA plans with a similar or better LCI as the clinical plan had an LSI that is 0.85 Gy larger than the LSI associated with the clinical plan, averaged over all patients (range 0.1–2.0 Gy, SD 0.6 Gy, $p < 0.001$).

Observer study

Table 2 shows the results of the observer study after presenting additional clinical patient data, totaling to 3 observers × 18 patients = 54 cases. Results before providing additional patient data are given in Supplementary Table 1. In all cases, the preferred plan was considered clinically acceptable. Furthermore, in all cases, one or more GOMEA plans were considered clinically acceptable. In 53/54 (98%) cases, a GOMEA plan was preferred. One observer selected a clinical plan once ($p = 0.173$), the two other observers never ($p = 0.038$). The coverage plan was preferred most often. Observers had different distributions of preferred plans, as shown in Supplementary Fig. 6. One observer checked all six dose distributions visually, while the other two first dismissed plans based on the DVIs, and only visually inspected the final one or two plans. The high-sparing plan was dismissed most often for insufficient coverage. For five patients (28%), one or more observers dismissed the clinical plan.

The preferred plan was changed eight times after presenting clinical information, of which five times by one observer. Half the changes were to increase coverage due to a high Gleason score, the other half to decrease coverage due to a low Gleason score or bad urinary flow. Five of these eight changes were made for patients who had a diffusion-weighted MRI.

For most of the plans, prostate coverage vs. urethra sparing was the dominating trade-off. In that case, observers generally looked for the plan with maximum prostate coverage (visually inspected or based on the $V_{100\%}^{prostate}$) while satisfying the urethra sparing criterion. Although, for some patients, prostate coverage of this plan was deemed insufficient, and a plan was chosen that violated the urethra sparing criterion. In the visual inspection of the dose distribution, observers focused on locations where the target was not covered. The whole prostate was considered to be target volume and an underdose can only be acceptable in parts where no tumor is expected. Observers also focused on hotspots and the activation of dwell positions in close proximity of the OARs. In a few cases, observers mentioned that they would like to try to manually improve GOMEA plans by disabling dwell positions or by spreading out the dose more evenly over multiple dwell positions. Overall, observers highly appreciated to see multiple plans simultaneously to get an impression of the achievable trade-offs. Decision-making was said to be harder when none of the presented plans satisfied all planning criteria (patients 4, 7, 14, and 18).

Three patients (5, 8, and 11) were discussed in the consensus meeting (Table 2). Observers easily came to an agreement. For Patient 11, the clinical plan was preferred over the other plans by one observer and later in the consensus meeting by all observers. For this patient, all three presented plans satisfy all planning criteria and observers agreed all plans are clinically acceptable. The coverage/sparing plan was dismissed based on the DVIs. Finally, the coverage plan was dismissed as well because of a high value of the additionally computed $D_{0.01cm3}^{rectum} = 186\%$ compared with $D_{0.01cm3}^{rectum} = 85\%$ of the clinical plan.

Table 2
Per patient, each observer indicated which of the six plans should be dismissed (−), which was the preferred plan after all clinical information was presented (+), and which plan was chosen in the consensus meeting for patients 5,8 and 11 (*).

	Presented plans in the observer study					
Patient	High Sparing	Sparing	Coverage/sparing	Coverage	High coverage	Clinical
1	−−−			++	+	
2	−−−		+	++		
3	−−		+	++	−	
4	−−−			++	+	
5	−−−	+		+*	+	
6	−−		++	+	−	
7	−		+++			−−
8	−−−	+		+	+*	
9	−−−		+	++		
10	−−−	+		++		
11	−−−		+	+		+*
12	−−−	+		++		
13	−−			++	+	−
14	++			+	−	−−
15	−			+++		−−
16	−−		++		+−	
17	−−−			+++		
18			+	+	+	−−−

For patients 5, 8, and 11, the resulting preferred plans after the consensus meeting are denoted by (*).

Discussion

The novel bi-objective treatment planning method that was clinically evaluated for the first time in this work automatically generates a large set of high-quality treatment plans, from which the planner can then select the desired plan. This makes treatment planning an insightful decision-making process instead of a trial-and-error optimization process.

Recordings of the clinical planning sessions showed that the bi-objective model corresponded well to our clinical practice. Not all planning criteria in our clinical protocol were included in the bi-objective model. First, $D_{90\%}^{prostate} > 100\%$ was omitted, which is a criterion that should be satisfied, but should not be maximized further. This distinction is important to notice and it highlights the importance of considering each planning criterion carefully when configuring the model. Second, $V_{150\%}^{prostate} < 50\%$ and $V_{200\%}^{prostate} < 20\%$ could be excluded in our case, as these criteria have loose aims that were always achieved. Alternatively, if these criteria would need to be included anyway by addition to the LSI, because a different clinical protocol is used, or when extending this method to a different treatment site, they could be rewritten as $D_{50\%}^{prostate} < 150\%$ and $D_{20\%}^{prostate} < 200\%$ to avoid comparing indices with different units (e.g., percentage of volume with percentage of dose).

GOMEA is by design able to handle more than two objectives. Additional criteria based on indices with different unit could therefore be added as third of fourth objective. Nevertheless, so far, for the BT treatment planning application, we limited ourselves to the use of two objectives. The main downside of additional objectives is that it would make it harder to visualize the trade-off curve, which potentially complicates the decision-making.

In the recordings, and during the visual inspection of the dose distributions in the observer study, observers focused not only on the planning criteria but on other aspects as well: hotspots, the activation of dwell positions near the OARs, and the location of areas where the target was not covered. Ensuring target coverage locally can be incorporated in optimization by, for example, indicating which subvolumes (e.g., quadrants) of the prostate should be fully covered because of tumor presence. The other aspects are not easily quantified. Earlier attempts have shown to potentially deteriorate plan quality, and different attempts were shown to be inconsistent (31).

For some plans, observers mentioned they would like to try to manually improve the plan, mainly to satisfy criteria that are not explicitly in the clinical protocol. By not allowing further improvements, it was showed that GOMEA plans are already clinically acceptable in their current form, without further tuning needed. Nevertheless, the potential to further improve the automatically generated plans by manual graphical optimization is of interest in future studies.

Several other planning approaches (4, 9) also model the treatment planning problem based on the (approximated) DVIs, but in a single-objective manner, where the trade-off is represented by weighting criteria. Fig. 3 however shows that the trade-off is patient-dependent and Supplementary Fig. 6 suggests that the preferred plan was in principle observer-dependent, although observers could come to an agreement in the consensus meeting. These confirm the value and validity of the DVI-based multi-objective planning approach in this work.

A limitation of the present study is that not for all patients the same set of clinical information was available to be presented in the observer study. The two-step approach in the observer study gives some insight into how this affected decision-making. Most changes of the preferred plan were made for patients who had a diffusion-weighted MRI available, which might suggest that the diffusion-weighted MRI is of additional value for decision-making. It also suggests that more changes would have been made when this information was available for all patients. However, with the small total number of changes observers made, it is unknown if having the same information for all patients would have led to consistent changes in plan preference.

Future work will be to evaluate applicability of the bi-objective optimization model to other planning criteria as used by other institutes. In addition, development of novel software tools that can allow fast navigation of GOMEA plans instead of preselecting five plans is of interest for the clinical implementation of our bi-objective treatment planning.

Conclusion

To conclude, we retrospectively evaluated a novel bi-objective BT planning method for use in our clinic. The analysis of current clinical planning sessions showed that the bi-objective model was easily configured and represented our clinical practice well. The observer study demonstrated that the bi-objective method automatically generates plans of high quality. For all patients and by all observers, resulting plans were preferred over the clinical plan in 98% of the cases. The ability to compare multiple high-quality plans was considered insightful and highly appreciated by the observers.

Acknowledgments

This work is part of the research program IPPSI-TA with project number 628.006.003, which is financed by the Netherlands Organisation for Scientific Research (NWO), Netherlands and Elekta, Sweden.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brachy.2018.12.010.

References

[1] Lessard E, Pouliot J. Inverse planning anatomy-based dose optimization for HDR-brachytherapy of the prostate using fast simulated annealing algorithm and dedicated objective function. *Med Phys* 2001; 28:773–779.

[2] Alterovitz R, Lessard E, Pouliot J, *et al*. Optimization of HDR brachytherapy dose distributions using linear programming with penalty costs. *Med Phys* 2006;33:4012–4019.

[3] Milickovic N, Lahanas M, Papagiannopoulo M, *et al*. Multiobjective anatomy-based dose optimization for HDR-brachytherapy with constraint free deterministic algorithms. *Phys Med Biol* 2002;47: 2263–2280.

[4] Morton GC, Sankreacha R, Halina P, *et al*. A comparison of anatomy-based inverse planning with simulated annealing and graphical optimization for high-dose-rate prostate brachytherapy. *Brachytherapy* 2008;7:12–16.

[5] Guthier CV, Damato AL, Viswanathan AN, *et al*. A fast multi-target inverse treatment planning strategy optimizing dosimetric measures for high-dose-rate (HDR) brachytherapy. *Med Phys* 2017;44:4452–4462.

[6] Kolkman-Deurloo IKK, Deleye XGJ, Jansen PP, *et al*. Anatomy based inverse planning in HDR prostate brachytherapy. *Radiother Oncol* 2004;73:73–77.

[7] Gorissen BL, den Hertog D, Hoffmann AL. Mixed integer programming improves comprehensibility and plan quality in inverse optimization of prostate HDR brachytherapy. *Phys Med Biol* 2013;58: 1041–1057.

[8] Holm Å, Larsson T, Tedgren ÅC. A linear programming model for optimizing HDR brachytherapy dose distributions with respect to mean dose in the DVH-tail. *Med Phys* 2013;40:81705.

[9] Siauw T, Cunha A, Atamtürk A, *et al*. IPIP: a new approach to inverse planning for HDR brachytherapy by directly optimizing dosimetric indices. *Med Phys* 2011;38:4045–4051.

[10] Deist TM, Gorissen BL. High-dose-rate prostate brachytherapy inverse planning on dose-volume criteria by simulated annealing. *Phys Med Biol* 2016;61:1155–1170.

[11] Dinkla AM, van der Laarse R, Kaljouw E, *et al*. A comparison of inverse optimization algorithms for HDR/PDR prostate brachytherapy treatment planning. *Brachytherapy* 2015;14:279–288.

[12] Maree SC, Bosman PAN, Niatsetski Y, *et al*. Improved class solutions for prostate brachytherapy planning via evolutionary machine learning. ESTRO 36. *Radiother Oncol* 2017;123:S96–S97.

[13] Lahanas M, Baltas D, Zamboglou N. A hybrid evolutionary algorithm for multi-objective anatomy-based dose optimization in high-dose-rate brachytherapy. *Phys Med Biol* 2003;48:399–415.

[14] Hoskin PJ, Colombo A, Henry A, *et al*. GEC/ESTRO recommendations on high dose rate afterloading brachytherapy for localised prostate cancer: an update. *Radiother Oncol* 2013;107:325–332.

[15] Luong NH, Bosman PAN. *Improving the performance of MO-RV-GOMEA on problems with many objectives using tchebycheff scalarizations*. Proc Genet Evol Comput Conf - GECCO '18; 2018. p. 705–712.

[16] Thieke C, Küfer KH, Monz M, *et al*. A new concept for interactive radiotherapy planning with multicriteria optimization: first clinical evaluation. *Radiother Oncol* 2007;85:292–298.

[17] Ruotsalainen H, Miettinen K, Palmgren J-E, *et al*. Interactive multi-objective optimization for anatomy-based three-dimensional HDR brachytherapy. *Phys Med Biol* 2010;55:4703–4719.

[18] Müller BS, Shih HA, Efstathiou JA, *et al*. Multicriteria plan optimization in the hands of physicians: a pilot study in prostate cancer and brain tumors. *Radiat Oncol* 2017;12:168.

[19] Craft DL, Hong TS, Shih HA, *et al*. Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy. *Int J Radiat Oncol Biol Phys* 2012;82:e83–e90.

[20] Morén B, Larsson T, Carlsson Tedgren Å. Mathematical optimization of high dose-rate brachytherapy - derivation of a linear penalty model from a dose-volume model. *Phys Med Biol* 2018;63:065011.

[21] Luong NH, Bouter A, Van Der Meer MC, *et al*. Efficient, effective, and insightful tackling of the high-dose-rate brachytherapy treatment planning problem for prostate cancer using evolutionary multi-objective Optimization Algorithms. In: Smith SL, Cagnoni S, Patton RM, editors. *Proc. Med. Appl. Genet. Evol. Comput. MedGEC Work. Genet. Evol. Comput. Conf. - GECCO-2017*. New York, New York: ACM Press; 2017. p. 1372–1379.

[22] Pieters BR, van der Grient JNB, Blank LECM, *et al*. Minimal displacement of novel self-anchoring catheters suitable for temporary prostate implants. *Radiother Oncol* 2006;80:69–72.

[23] Luong NH, Alderliesten T, Bel A, *et al*. Application and benchmarking of multi-objective evolutionary algorithms on high-dose-rate brachytherapy planning for prostate cancer treatment. *Swarm Evol Comput* 2018;40:37–52.

[24] Maree SC, Kooreman ES, Luong NH, *et al*. PO-1020: better plans and easy plan selection via bi-objective optimization for HDR prostate brachytherapy. In: *Proc. Eur. Soc. Radiother. Oncol. Conf. - ESTRO-2018*;. p. S571–S572.

[25] Prada PJ, Cardenal J, Blanco AG, *et al*. High-dose-rate interstitial brachytherapy as monotherapy in one fraction for the treatment of favorable stage prostate cancer: toxicity and long-term biochemical resultsProstate HDR brachytherapy as monotherapy in one fraction. *Radiother Oncol* 2016;119:411–416.

[26] Rivard MJ, Coursey BM, DeWerd LA, *et al*. Update of AAPM Task Group No. 43 Report: a revised AAPM protocol for brachytherapy dose calculations. *Med Phys* 2004;31:633–674.

[27] Van der Meer MC, Bosman PAN, Pieters BR, *et al*. EP-2236: Sensitivity of Dose-Volume Indices to Organ Reconstruction Settings in HDR Prostate Brachytherapy. In: *Radiother. Oncol. Elsevier Masson SAS*;. p. S1236.

[28] Bouter A, Luong NH, Witteveen C, *et al*. The multi-objective real-valued gene-pool optimal mixing evolutionary algorithm. In: *Proc. Genet. Evol. Comput. Conf. - GECCO '17*;. p. 537–544.

[29] Luong NH, Alderliesten T, Pieters BR, *et al*. PV-0256: fast and insightful bi-objective HDR prostate brachytherapy planning. In: *Proc. Eur. Soc. Radiother. Oncol. Conf. - ESTRO-2018*;. p. S130.

[30] Bouter A, Alderliesten T, Bel A, *et al*. *Large-scale parallelization of partial evaluations in evolutionary algorithms for real-world problems*.Proc. Genet. Evol. Comput. Conf. - GECCO-2018; 2018. p. 1199–1206.

[31] Balvert M, Gorissen BL, den Hertog D, *et al*. Dwell time modulation restrictions do not necessarily improve treatment plan quality for prostate HDR brachytherapy. *Phys Med Biol* 2015;60:537–548.