Radiotherapy and Oncology 158 (2021) 253–261

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Pre-clinical validation of a novel system for fully-automated treatment planning

Rik Bijman *, Abdul Wahab Sharfo, Linda Rossi, Sebastiaan Breedveld, Ben Heijmen

Department of Radiotherapy, Erasmus MC Cancer Institute, University Medical Center Rotterdam, The Netherlands; Department of Radiotherapy, Erasmus MC Cancer Institute, GD Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received 16 November 2020
Received in revised form 24 February 2021
Accepted 2 March 2021
Available online 10 March 2021

Keywords:
Automated treatment planning
Multi-criterial optimization (MCO)
Validation

ABSTRACT

Introduction: Many approaches for automated treatment planning (autoplanning) have been proposed and investigated. Autoplanning can enhance plan quality compared to 'manual' trial-and-error planning, and decrease routine planning workload. A few approaches have been implemented in commercial treatment planning systems (TPSs). We performed a pre-clinical validation of a new system ('NovelATP') that is based on fully-automated multi-criterial optimization (MCO). The aim of NovelATP is to automatically generate for each patient a single high-quality, Pareto-optimal plan without manual Pareto navigation.
Material and methods: Validation was performed by generating VMAT/IMRT plans for conventional treatment of prostate cancer (101 pts), prostate SBRT (20 pts), bilateral head-and-neck cancer (50 pts) and rectal cancer treated at an MR-Linac (23 pts). NovelATP autoplans were compared to plans that were generated with our in-house autoplanning system. In many previous validation studies, the latter system consistently showed enhanced plan quality when compared to manual planning.
Results: Dosimetrical differences between NovelATP and benchmark plans were on average small and presumably not clinically relevant, pointing at high NovelATP dosimetric plan quality. MUs were 11–19% higher with NovelATP. NovelATP delivery times were up to 12% longer. Overall, there was a slight disadvantage for NovelATP regarding gamma analyses. Calculation times for NovelATP plans were between 29 and 151 min with no overall differences with the benchmark plans.
Conclusion: The new autoplanning system was able to produce high-quality plans for four highly different planning protocols/treatment sites with a total of 194 patients investigated.

In the past decades, significant advances in radiotherapy delivery have been made to adequately irradiate tumors while at the same time better spare the surrounding organs at risk (OARs) [1–5]. Adequate treatment planning is a pre-requisite for application of the latest treatment approaches to their full extent. For years, treatment planning has been an interactive trial-and-error process performed by planners that try for each patient to steer the treatment planning system (TPS) towards generation of an acceptable, high-quality plan ('manual planning'). It is well-known that plan quality in manual planning may be sub-optimal [6–11]. Suboptimal plans can seriously impact treatment outcome [12,13].

In recent years, many solutions for automated treatment planning (autoplanning) have been proposed to enhance quality and reduce planning workload, as summarized in the review by Hussein et al. [11]. Two general-purpose autoplanning approaches have been implemented in commercial TPSs and extensively tested, based on knowledge-based planning [14–16], and protocol-based automatic iterative optimization [17–19]. A dedicated application for breast autoplanning was also implemented in a commercial TPS [20,21]. Scripting in commercial TPS has also resulted in several clinical autoplanning applications [22,23]. Autoplanning has also been commercially implemented for so-called *a posteriori* multi-criterial optimization (MCO): automated generation of a Pareto front, followed by final plan selection by a planner using Pareto navigation [24–27].

At our center, a system has been developed for fully-automated *a priori* MCO, generating a single Pareto-optimal plan per patient. For each planning protocol, the system is *a priori* configured to ensure that the generated Pareto-optimal plans are also clinically favorable. The system consists of the in-house Erasmus-iCycle optimizer, which is coupled to the commercial Monaco TPS (Elekta AB, Stockholm, Sweden) for final plan generation [10,11,28]. Several studies have shown enhanced plan quality with this

* Corresponding author at: Department of Radiotherapy, Erasmus MC Cancer Institute, University Medical Center Rotterdam, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands.
E-mail addresses: r.bijman@erasmusmc.nl (R. Bijman), a.sharfo@erasmusmc.nl (A.W. Sharfo), l.rossi@erasmusmc.nl (L. Rossi), s.breedveld@ersasmusmc.nl (S. Breedveld), b.heijmen@erasmusmc.nl (B. Heijmen).

https://doi.org/10.1016/j.radonc.2021.03.003
0167-8140/

'ErasmusATP' (Erasmus Automated Treatment Planning) platform, compared to manual, trial-and-error planning [29–36]. ErasmusATP is in clinical use at Erasmus MC since 2012.

Recently, a TPS vendor (Elekta AB) used the concept of *a priori* MCO to develop a new autoplanning application ('NovelATP'). The aim of this study was to perform a comprehensive pre-clinical validation of NovelATP. To this purpose, the ErasmusATP and NovelATP platforms were both configured for four planning protocols, followed by pairwise plan comparisons. Plan deliverability was checked by dosimetric measurements at a linac.

Material and methods

ErasmusATP

ErasmusATP has been described in detail in the literature [10,28]. Here a brief summary is provided. Some details can also be found in Electronic Appendix 1.

The ErasmusATP workflow, used for validation of NovelATP, is schematically depicted in Fig. 1a. Input for the applied two-step approach for fully-automated plan generation is a contoured planning CT and a planning protocol specific planning wish-list, introduced in more detail below. Based on this input, a fluence map optimization (FMO) is first performed using the Erasmus-iCycle multi-criterial optimizer [28]. The obtained Pareto-optimal dose distribution is then used to construct a patient-specific Monaco template (see below), which is subsequently used for final plan generation with the Monaco TPS (Elekta AB, Stockholm, Sweden). As for any plan generation with Monaco, generation of this final plan starts with an FMO (not to be confused with the FMO in Erasmus-iCycle, see above), followed by an optimization of segments ('Segmentation' in Fig. 1a). Due to the applied *patient-specific* template in Monaco, the final plan can be generated without any human interference, while mimicking the Erasmus-iCycle dose distribution and converting it into a deliverable plan. With this two-step approach, the final plan is generated with a CE-marked TPS, allowing clinical use.

The planning protocol specific wish-lists instruct Erasmus-iCycle how to perform for each individual patient the automated multi-criterial plan generation [28]. A wish-list is defined in terms of (1) hard planning constraints, and (2) objective functions with goal values, ordered in importance by assigned priorities. Electronic Appendix A1 provides a description of wish-list creation (configuration), and its use in Erasmus-iCycle for prioritized plan generation.

The patient-specific Monaco templates (Fig. 1a) each contain a set of Monaco cost functions with corresponding goal values. For each planning protocol, the template cost functions are fixed for the involved patient group, but goal values are patient-specific as they are derived from obtained Erasmus-iCycle doses (Fig. 1a). Applied Monaco templates account for intrinsic differences between Erasmus-iCycle and Monaco, e.g. in available cost functions, dose calculation engine, and optimizer.

NovelATP

The NovelATP workflow is schematically depicted in Fig. 1b. As in ErasmusATP, autoplanning starts with wish-list driven multi-criterial FMO. However, in contrast to ErasmusATP, a patient-

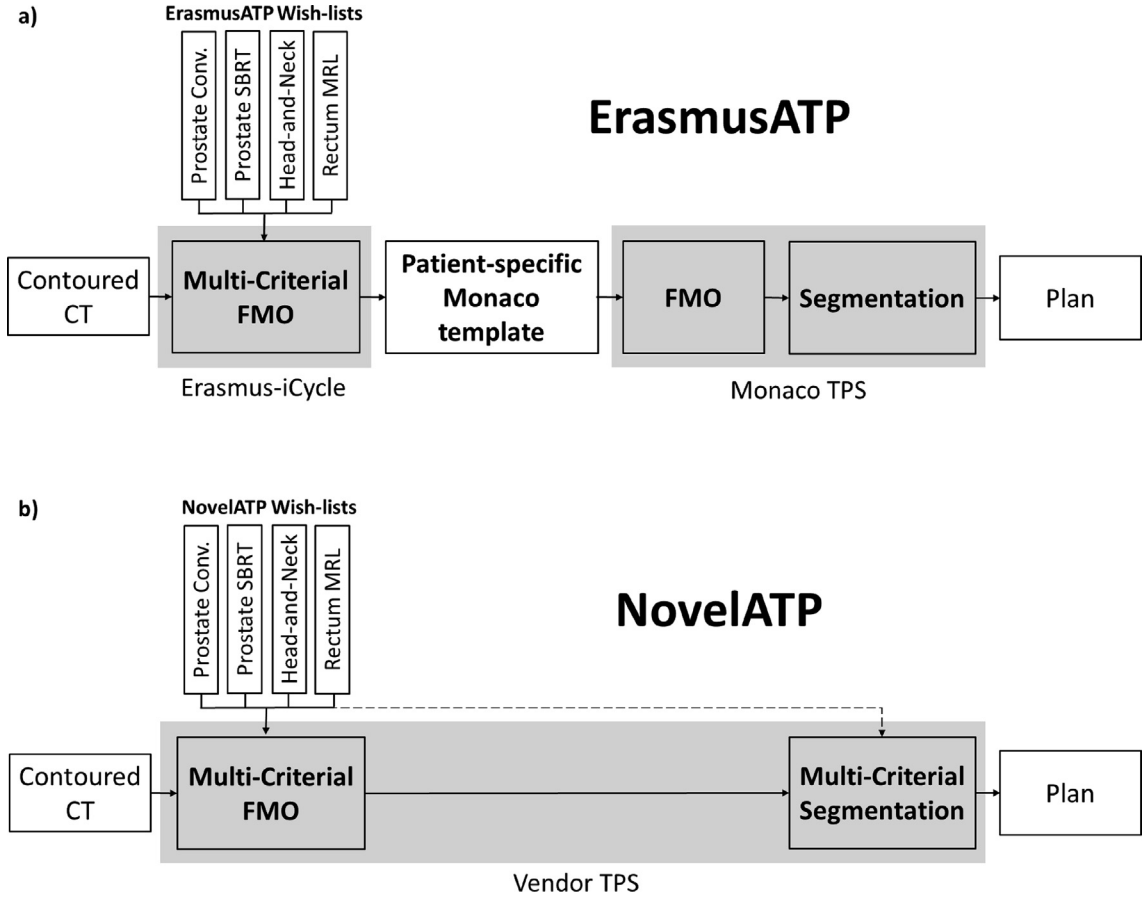

Fig. 1. Schematic overview of (a) ErasmusATP and (b) NovelATP, depicting the most important similarities and differences between the two systems. See text for further explanations.

specific Monaco template and a second FMO are no longer needed to obtain a deliverable plan (Compare Fig. 1a and b). Instead, the FMO dose distribution is input for a multi-criterial optimization of MLC segments, using a weighted-sum cost function with constraint and objective functions as used in FMO and defined by the wish-list. Segment weights and shapes are iteratively adapted by minimizing the cost function values. During segmentation, two types of constraints are used: 'clinical' constraints (e.g. maximum cord dose) that are assigned a higher initial and maximum weight than so-called 'planning' constraints (e.g. introduced in the wish-list to enhance dose conformity), and objectives. The latter get maximum weights in accordance to their priority in the wish-list.

Main technical differences between NovelATP and ErasmusATP

Although the multi-criterial FMO in ErasmusATP and NovelATP are both wish-list driven and conceptually similar, implementation differences are large. NovelATP FMO is based on completely new code, using a different mathematical solver and patient model. Moreover, the available cost functions in wish-lists differ from those in ErasmusATP (see Electronic Appendix A2 for details). As explained in the above paragraph, approaches for segmentation of FMO plans are also completely different: in NovelATP, both FMO and segmentation are wish-list driven, while in ErasmusATP the wish-list is only used in FMO (compare Fig. 1a and b.).

Comparison of NovelATP with ErasmusATP

ErasmusATP and NovelATP were first configured for the investigated planning protocols (below), aiming at high(est)-quality plans according to the clinical planning tradition, while using the same beam energy, delivery approach (VMAT or IMRT) and segmentation settings (Electronic Appendix A2 for details). As explained above, configuration of ErasmusATP for a planning protocol entails creation of an appropriate wish-list and a basic Monaco template. For NovelATP, only a wish-list needs to be created. After the configurations, final plans were generated for all study patients. There was no manual fine-tuning of autoplans. PTV and OAR plan parameters used for clinical plan evaluations, patient hot spots, conformality and dose bath for the study patients were then pairwise compared. Two-sided Wilcoxon signed-rank tests were used to assess statistical significance of observed differences ($p < 0.05$).

For 10 randomly selected patients, ErasmusATP and NovelATP plans were delivered at a linac to assess delivery accuracy and differences in delivered MUs and delivery time. For patient selection, all study patients were put on a list in arbitrary order, followed by a random selection of 10. The dosimetric QA took place on two consecutive days using a PTW 2D-Array seven29™ and Octavius™ phantom (PTW, Freiburg, Germany) were used. Comparisons with TPS predictions were performed with VeriSoft version 7 (PTW, Freiburg, Germany) with 5% cut-off, 3% global maximum dose and 3 mm distance-to-agreement criteria used in the γ-analyses, following our clinical protocol and in line with the AAPM Task Group 119 report [37,38] A γ-passing criterion of 90% and mean γ-criterion of 0.5 were used, as clinically applied. Plan optimization times were compared, running plans for a fixed group of patients on identical hardware (dual Intel Xeon E2690). Timing information was obtained from log files.

Patients and planning protocols

NovelATP was validated for four planning protocols:

Conventional radiotherapy for prostate cancer ('Prostate Conventional'): All 101 prostate cancer patients, treated at Erasmus MC with the investigated protocol between February and June 2019 were included. Treatment depended on risk of seminal vesicle involvement [39]: (i) planning target volume (PTV) consisting of prostate + 0.5 cm isotropic margin, except for 0.7 cm inferiorly (PTV_1), treated with 60 Gy, delivered in 20 fractions (10 patients), (ii) PTV_1 and PTV_2 (PTV_2 = PTV_1 + seminal vesicles with 0.8 cm isotropic margin), treated with 60 Gy and 56 Gy, respectively, delivered in 20 fractions (53 patients), (iii) PTV_2 treated with 60 Gy in 20 fractions (38 patients). 10 MV single-arc VMAT was used for (i) and (iii), and dual-arc for (ii).

The aim was to deliver 95% of the prescribed doses to 99% of the PTVs ($V_{95\%}$=99%), while keeping the high dose PTV $V_{107\%} \leq 0.3\%$, and PTV $D_{max} \leq 110\%$. OAR hard constraints were defined for anus and femoral heads. Within the constraints, reduction of high rectum doses was the most important OAR planning aim, followed by lowering the mean dose to the rectum, anus, bladder, and maximum doses in the femoral heads.

Average PTV_1 and PTV_2 volumes were 137 cc (range: 73–280 cc) and 186 cc (range: 136–304 cc), respectively. All patients had N and M stage 0, while 14%, 34%, 51% and 1% of the patients were staged T1, T2, T3 and T4, respectively.

Stereotactic body radiation therapy for prostate cancer ('Prostate SBRT'): The PTV of the 20 consecutive patients consisted of the prostate + 0.3 cm isotropic margin. The prescribed dose was 38 Gy, delivered in four fractions using 10 MV dual-arc VMAT [40].

Dose distributions were highly heterogeneous aiming at high doses in the peripheral zone, while sparing the urethra [40]. The goal for the PTV was to obtain a $V_{100\%}$=95%. The minimum prostate dose should be >89.5% of the prescribed dose, i.e. 34 Gy. OAR hard constraints were defined for rectum, bladder, urethra, rectum mucosa and femoral heads. Within imposed constraints, reduction of (especially high) rectum dose was the important OAR planning objective. Subsequently, bladder high doses, rectum and bladder mean doses, urethra mean, and high doses and femoral heads maximum doses were reduced as much as possible.

Average PTV volume was 93 cc (range: 61–142 cc). All patients had N and M stage 0, while 58% and 42% of the patients were staged T1 and T2, respectively.

Conventional radiotherapy for bilateral head-and-neck cancer ('Head-and-Neck'): Fifty locally advanced head-and-neck cancer patients (19 oropharynx, 14 larynx, 17 hypopharynx), consecutively treated between February and June 2019 were included. Prescribed doses were 70 Gy to the primary tumor and pathological lymph nodes, and 54.25 Gy to the elective nodal areas, simultaneously delivered in 35 fractions. $PTV_{54.25Gy}$ consisted of the primary CTVs and the elective lymph node CTVs, expanded by 0.5 cm, while clipping at the patient surface by 0.5 cm. PTV_{70Gy} consisted of the primary CTVs, each expanded with a 0.5 cm margin, again clipped at the patient surface by 0.5 cm. In line with our clinical practice, for optimal coverage of the superficial targets, a flash margin (0.5–1 cm) was used (Monaco TPS feature "auto flash" that creates a margin containing voxels that extend beyond the surface of the patient into the air. This opens the leaves to conform to a virtual target). Dual-arc VMAT with 6MV photons was used.

For both PTVs, the aim was to obtain a coverage $V_{95\%} \geq 98\%$ with $V_{107\%} \leq 2$cc, and PTV $D_{mean} \leq 1.02\%$ and $D_{max} \leq 110\%$, for CTVs $V_{95\%}$ = 100%. Hard planning constraints were defined for spinal cord, brainstem and cochleas. OAR planning goals were, in order of priority, maximum reduction of mean doses in parotid glands, submandibular glands (SMGs), swallowing muscles (i.e. MCS, MCP, MCI, MCM), oral cavity, larynx, esophagus and cochlea's, and the maximum doses to the brainstem and spinal cord.

Average PTV_1 and PTV_2 volumes were 243 cc (range: 30–722 cc) and 678 cc (range: 185–1147 cc), respectively. All patients had M stage 0, and 2% were T1N1; 48%, 2% and 12% were T2N0, T2N1, and T2N2, respectively; 16% and 10% T3N0 and T3N2, respectively; and 2%, 2% and 6% T4N0, T4N1 and T4N2, respectively.

MR-Linac radiotherapy for rectum cancer ('Rectum MRL'): Data of the first 23 patients treated on a Unity MR-Linac (Elekta AB, Stockholm, Sweden) at the Netherlands Cancer Institute was used [36]. The clinical target volume (CTV) consisted of the GTV, expanded isotropically with a margin of 1 cm for subclinical disease, plus regional lymph nodes (mesorectal, iliac, and depending on GTV location and N-stage, obturator and/or presacral) with a 0.5 cm margin. The PTV was constructed by an anisotropic expansion of the CTV with a margin of 1 cm in all directions except for 1.5 cm anterior to the mesorectal region. 50 Gy was delivered in fractions of 2 Gy, using 9-beam IMRT [36].

For the PTV, planning goals were $V_{95\%} \geq 99\%$, $V_{107\%} < 1\%$ and $D_{mean} > 50$ Gy. The main OAR planning aim was reduction of mean bladder and bowel doses. For optimization purposes, bladder and bowel were combined into one composite OAR [36].

Average PTV volume was 1126 cc (range: 780–1530 cc). All patients had M stage 0, and 7% and 9% were T2N0, T2N1, respectively; 7%, 26% and 26% T3N0, T3N1 and T3N2, respectively; and 4%, 17% and 4% T4N0 T4N1 and T4N2, respectively.

For all treatment sites local contouring guidelines were followed for OAR definitions. High, intermediate and low dose conformality were compared based on the conformity index (CI (x) = volume receiving × Gy or x% of prescription dose/PTV volume). For prostate SBRT, V10Gy, V20Gy and V30Gy were also evaluated.

Results

All wish-lists for ErasmusATP and NovelATP are provided in Electronic Appendix A2.

Prior to plan comparisons, all ErasmusATP and NovelATP plans were rescaled to the clinically desired PTV coverages (above) to avoid bias in analyses of OAR doses. Prior to the re-scaling, the ErasmusATP/NovelATP plan coverages were 98.9% ± 0.6%/99.1% ± 0.3% for Prostate Conventional, 95.3% ± 0.2%/95.2% ± 0.7% for Prostate SBRT, 99.4% ± 0.5%/98.8% ± 0.6%, for Head-and-Neck and 99.4% ± 0.4%/99.1% ± 0.4% for Rectum MRL, and therefore only minor re-scaling was needed.

As summarized in Figs. 2 and 3 and Tables A3.1-A3.4 in Electronic Appendix A3, median/mean dosimetric differences between NovelATP and ErasmusATP were generally small and probably clinically irrelevant, although for dosimetrical parameters often statistically significant, sometimes in favor of NovelATP, other times in favor of ErasmusATP. At an individual patient level, differences may sometimes be clinically significant. Dose distributions of example patients are presented in Fig. 4.

A general trend for Prostate Conventional, Prostate SBRT and Head-and-Neck cancer was high similarity between the two autoplanning systems in obtained doses for the highest wish-list priorities. Solely for Rectum MRL, the highest prioritized objective, i.e. the mean dose in the composite OAR, showed a small benefit for NovelATP, however this was at a trade-off with higher doses in this OAR (Fig. 2).

Across all treatment groups it was generally observed that median and low doses were similar or slightly lower with NovelATP, and high doses were lower with ErasmusATP (Fig. 2).

NovelATP resulted in slightly better low dose conformality for Prostate SBRT and Rectum MRL, while for Prostate Conventional and Head-and-Neck ErasmusATP was slightly better (Figs. 3 and 4). High dose conformality was similar for the two systems.

For all treatment sites, plans generated with NovelATP had generally larger amounts of MUs and longer delivery times (Table 1), with the same segmentation settings (Electronic Appendix 2). Numbers of segments were generally comparable, except for Prostate SBRT, for which significantly more segments were used with ErasmusATP. All plans fulfilled the 90% γ-passing rate criterion and only one of the plans (NovelATP) had a mean$\gamma > 0.5$. Calculation times with NovelATP were between 29 and 151 minutes, depending on protocol and case. NovelATP planning took significantly shorter for Prostate Conventional and Head-and-Neck and longer for Prostate SBRT and Rectum MRL (Table 1).

Discussion

This paper reports on a pre-clinical validation of a new system for fully-automated multi-criterial optimization (MCO) of treatment plans. In contrast to Pareto navigation based MCO, only a single, final plan is automatically generated for each patient. The system needs to be configured for each planning protocol to ensure that the generated plans are clinically favorable. This configuration is then used for the whole patient cohort, without patient-specific modification.

For validation, plans generated with the new autoplanning system were compared to plans generated with our in-house clinical autoplanning workflow. Both systems use wish-list driven prioritized optimization, but both the implementation and use of the wish-list differ substantially (M&M section).

In the literature, validations of autoplanning systems were generally performed by comparison of autoplans with competing manually generated plans [10,15,18]. Such validations are limited by the well-known problems with quality and consistency of manually generated plans [6,7]. In many previously published institutional and international studies, it was demonstrated that we could configure our in-house autoplanning system for superior dosimetric plan quality, compared to manually generated plans [29–36]. To provide in this study the 'best possible' competing autoplans as benchmark for validation of the new autoplanning system, maximum attention was paid to the optimal configuration of the in-house system. To this purpose, all in-house wish-lists were generated by researchers with vast configuration experience, and knowledge of the investigated planning protocols. For each of the four investigated planning protocols, our expert planner for the protocol was in the lead of generation of a configuration that maximally complied with all institutional planning aims. To minimize bias towards one of the systems and to maintain consistency in system configuration, for each protocol, the same investigator was in the lead of the configuration of the new autoplanning system. For both systems, development of the configurations was frequently discussed in the full research group with a focus on maximizing plan quality. There were no time restrictions for configurations. Nevertheless, optimal wish-list configuration cannot be guaranteed. A way to further investigate quality of the applied wish-lists could be Pareto-navigation studies, with for each patient the generated autoplan as a starting point for the navigation.

Dosimetric plan parameters and DVHs obtained with the new autoplanning system were on average similar to those produced by the benchmark, with sometimes relevant differences at an individual patient level; depending on the patient, this could be in favour of the new autoplanning system or the in-house system. The required MUs were slightly higher for the new system (mean differences between 11% and 19%, depending on the planning protocol). For prostate SBRT, the number of segments was on average 12% lower with the new system. For three protocols, delivery times were enhanced with the new system (mean differences: 6%, 9% and 12%). For two protocols, there was a slight advantage for the in-house system regarding γ-analyses. This could possibly be due to large differences between the two autoplanning systems in segmentation of FMO plans (M&M section). Calculation times for the

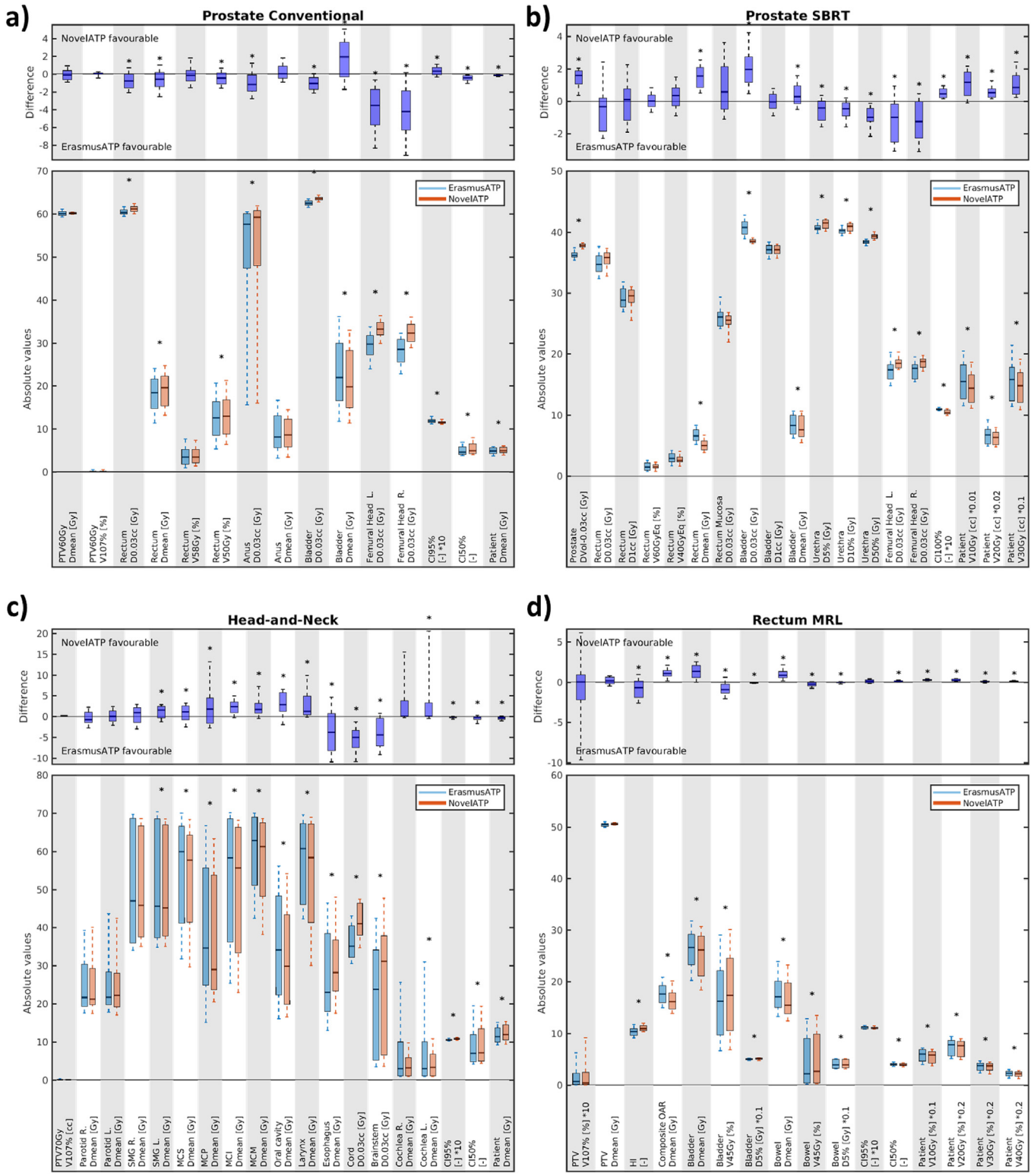

Fig. 2. Boxplots showing distributions of parameter value differences (top) and absolute parameter values (bottom) for ErasmusATP and NovelATP. The boxes represent the 25th to 75th percentile of the data with the median depicted by the horizontal line. The whiskers show the 10th to 90th percentile. Statistically significant differences between ErasmusATP and NovelATP are indicated by asterisks. SMG = submandibular gland, MCS = superior constrictor muscle, MCM = middle constrictor muscle, MCI = inferior constrictor muscle, MCP = palatopharynqeus muscle.

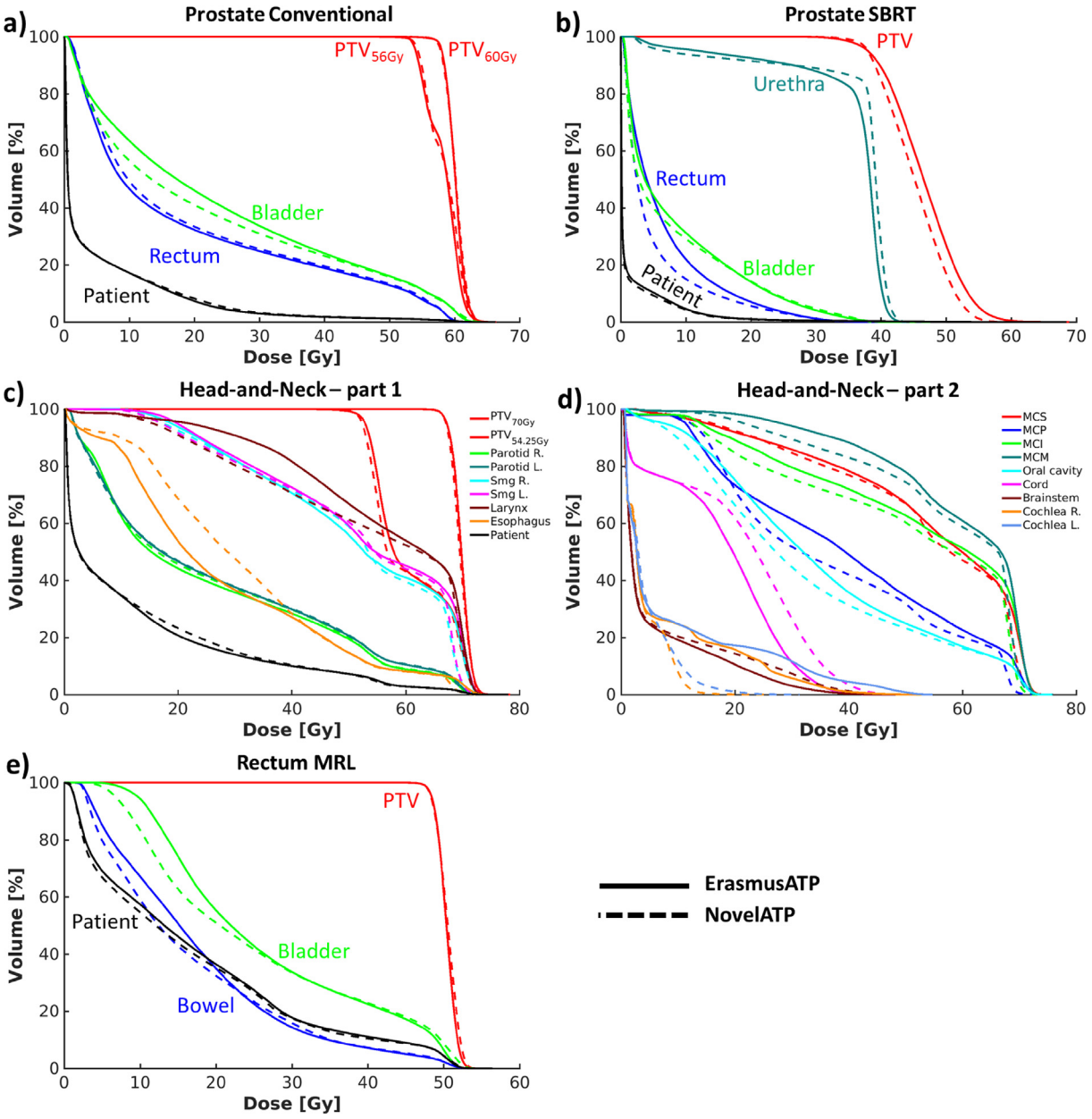

Fig. 3. Population mean DVHs for the four investigated planning protocols. SMG = submandibular gland, MCS = superior constrictor muscle, MCM = middle constrictor muscle, MCI = inferior constrictor muscle, MCP = palatopharynqeus muscle.

new autoplanning system were between 29 and 151 minutes, depending on protocol and case, which was sometimes shorter and sometimes longer than for the in-house system.

The aim of this study was to create 'best possible' plans with the new autoplanning system and compare them with 'best possible' plans, generated with the in-house benchmark autoplanning system. As the two systems showed similar overall dosimetric autoplan quality, we have indirectly demonstrated that the new autoplanning system has a high potential to improve quality compared to manual plans, as previously also observed for the benchmark system (see above). Of course, this can depend on many factors, such as quality of the manual plans, type of planning protocol, etc. It will certainly also depend on the quality of the configuration of the new system, which is largely dependent on the quality of the wish-list (See M&M). In this study, configuration was performed by highly experienced researchers. More studies are needed in centers with less experience to get a more complete picture of the potential of the novel system, and to understand more about potential re-use of prior wish-lists for new wish-list development and the learning curve for using it to its full potential. These investigations will then naturally also include different planning protocols and patient groups.

The RATING guidelines for treatment planning studies [41] were used for preparing the manuscript. Two authors (RB, LR)

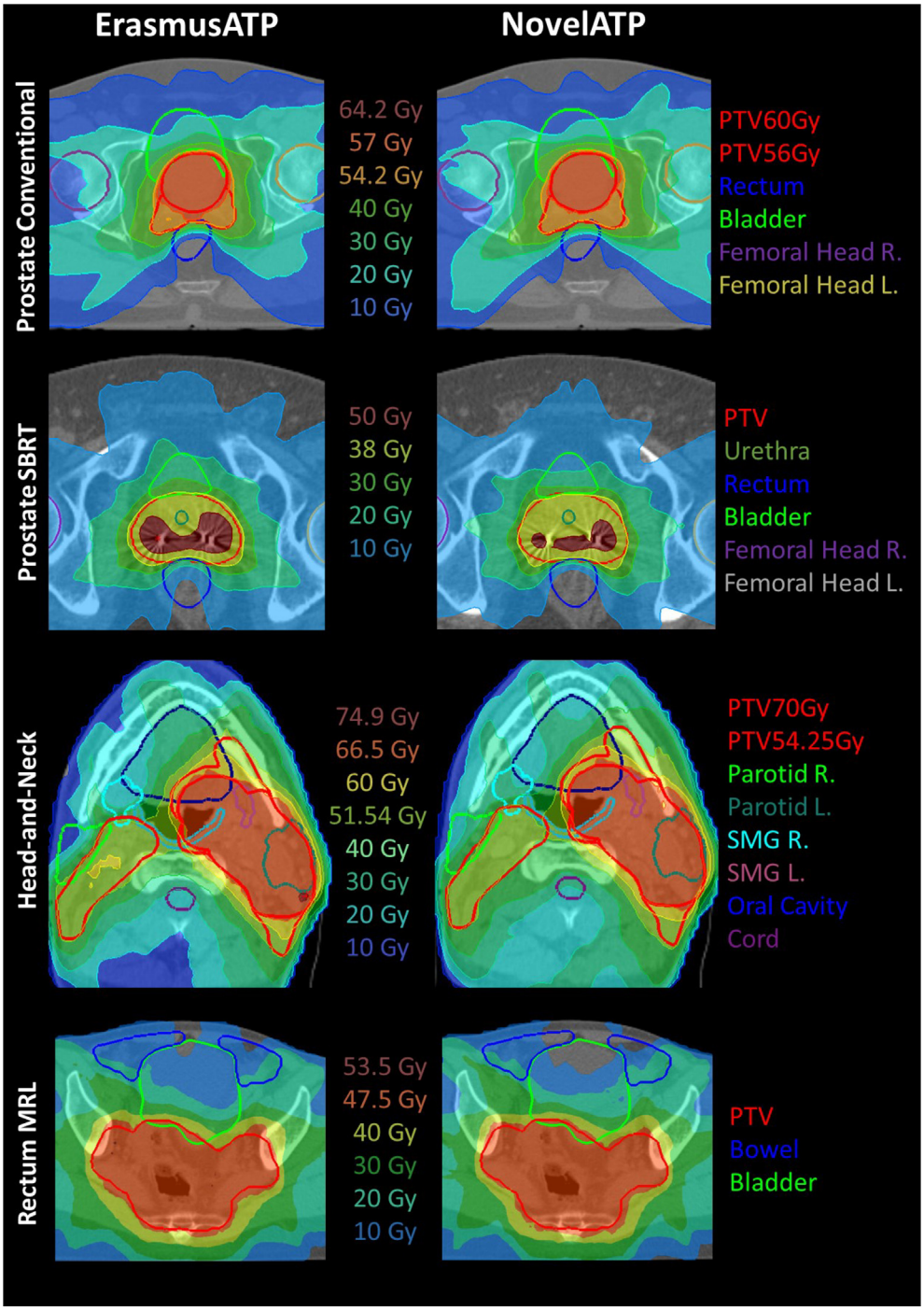

Fig. 4. Dose distributions for an example patient of each planning protocol. Left: ErasmusATP, Right: NovelATP. Each example patient was selected as being the patient with the median difference in the highest prioritized OAR objective: rectum near maximum dose for Prostate Conventional, rectum near maximum dose for Prostate SBRT, parotid mean dose for Head-and-Neck and composite OAR mean dose for Rectum MRL. SMG = submandibular gland.

independently arrived at RATING scores of 79% and 83%, respectively.

In conclusion, a new autoplanning system was validated by benchmarking it with another, well-established autoplanning system. The new system was able to produce high-quality plans for conventional prostate cancer treatment, prostate SBRT, head-and-neck cancer, and rectum cancer treated at an MR-Linac, for in total 194 investigated patients.

Table 1
Non-dosimetric plan comparisons for a subset of 10 patients per planning protocol. Mean values with ranges are reported. Except for the γpassrate, negative differences point at an advantage for NovelATP. Linac measurements were not performed for Rectum MRL. Difference of the averages is reported (Difference [%]).

		Prostate Conventional	Prostate SBRT	Head-and-Neck	Rectum MRL
MUs	*ErasmusATP [#]*	871 [758–959]	4328 [3776–5297]	767 [667–951]	645 [558–829]
	NovelATP [#]	970 [872–1037]	5148 [4188–5711]	866 [729–1008]	770 [657–918]
	Difference [%]	11.4	19.0	12.9	19.4
	p-value	0.006	0.004	0.002	0.002
Segments	*ErasmusATP [#]*	125 [115–132]	256 [244–269]	216 [202–228]	75 [56–97]
	NovelATP [#]	121 [117–124]	225 [209–239]	212 [194–230]	72 [59–84]
	Difference [%]	−2.9	−12.1	−1.9	−3.5
	p-value	0.1	0.002	0.2	0.3
Delivery time	*ErasmusATP [s]*	100 [89–109]	386 [346–468]	166 [149–187]	298 [226–413]
	NovelATP [s]	105 [96–113]	461 [389–513]	185 [168–210]	311 [252–366]
	Difference [%]	5.9	19.5	11.8	4.3
	p-value	0.03	0.002	0.002	0.2
QA – γ passrate	*ErasmusATP [%]*	99.2 [97.3–100]	99.9 [99.3–100]	98.4 [96.7–99.4]	–
	NovelATP [%]	97.7 [92.9–100]	98.1 [94.9–100]	97.8 [91.8–99.4]	–
	Difference [%]	−1.4	−1.8	−0.6	–
	p-value	0.004	0.02	0.9	–
QA – mean γ	*ErasmusATP [-]*	0.35 [0.29–0.45]	0.37 [0.30–0.33]	0.34 [0.30–0.40]	–
	NovelATP [-]	0.39 [0.31–0.46]	0.43 [0.33–0.51]	0.36 [0.30–0.49]	–
	Difference [%]	11.0	16.0	4.1	–
	p-value	0.08	0.004	0.6	–
Planning Time	*ErasmusATP [min]*	65 [52–79]	52 [35–69]	142 [85–219]	25 [20–33]
	NovelATP [min]	45 [35–57]	57 [48–69]	108 [65–151]	34 [29–44]
	Difference [%]	−31.1	9.0	−24.0	33.3
	p-value	0.002	0.3	0.002	0.004

Conflicts of interest

This work was in part funded by a research grant of Elekta AB (Stockholm, Sweden). Erasmus MC Cancer Institute also has a research collaboration with Accuray Inc, Sunnyvale, USA. No additional external funding was received for this study. The funders had no role in study design, data collection and analysis.

Acknowledgements

The authors want to thank Elekta AB for NovelATP training and assistance. Erik Loef (Erasmus MC) for his assistance on performing the QA measurements. The NKI in Amsterdam is acknowledged for providing MRL data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2021.03.003.

References

[1] Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. Lancet Oncol 2011;12:127–36.
[2] Mukesh MB, Barnett GC, Wilkinson JS, Moody AM, Wilson C, Dorling L, et al. Randomized controlled trial of intensity-modulated radiotherapy for early breast cancer: 5-year results confirm superior overall cosmesis. J Clin Oncol 2013;31:4488–95.
[3] Chun SG, Hu C, Choy H, Komaki RU, Timmerman RD, Schild SE, et al. Impact of intensity-modulated radiation therapy technique for locally advanced non-small-cell lung cancer: a secondary analysis of the NRG oncology RTOG 0617 randomized clinical trial. J Clin Oncol 2017;35:56–62.
[4] Viani GA, Viana BS, Martin JEC, Rossi BT, Zuliani G, Stefano EJ. Intensity-modulated radiotherapy reduces toxicity with similar biochemical control compared with 3-dimensional conformal radiotherapy for prostate cancer: a randomized clinical trial. Cancer 2016;122:2004–11.
[5] Staffurth J, Radiotherapy DB. A review of the clinical evidence for intensity-modulated radiotherapy. Clin Oncol (R Coll Radiol) 2010;22:643–57.
[6] Nelms BE, Robinson G, Markham J, Velasco K, Boyd S, Narayan S, et al. Variation in external beam treatment plan quality: an inter-institutional study of planners and planning systems. Pract Radiat Oncol 2012;2:296–305.
[7] Berry SL, Boczkowski A, Ma R, Mechalakos J, Hunt M. Interobserver variability in radiation therapy plan output: results of a single-institution study. Pract Radiat Oncol 2016;6:442–9.
[8] Hansen CR, Nielsen M, Bertelsen AS, Hazell I, Holtved E, Zukauskaite R, et al. Automatic treatment planning facilitates fast generation of high-quality treatment plans for esophageal cancer. Acta Oncol 2017;56:1495–500.
[9] Rossi L, Sharfo AW, Aluwini S, Dirkx M, Breedveld S, Heijmen B. First fully automated planning solution for robotic radiosurgery - comparison with automatically planned volumetric arc therapy for prostate cancer. Acta Oncol 2018;57:1490–8.
[10] Heijmen B, Voet P, Fransen D, Penninkhof J, Milder M, Akhiat H, et al. Fully automated, multi-criterial planning for Volumetric Modulated Arc Therapy - an international multi-center validation for prostate cancer. Radiother Oncol 2018;128:343–8.
[11] Hussein M, Heijmen BJM, Verellen D, Nisbet A. Automation in intensity modulated radiotherapy treatment planning-a review of recent innovations. Br J Radiol 2018;91:20180270.
[12] Peters LJ, O'Sullivan B, Giralt J, Fitzgerald TJ, Trotti A, Bernier J, et al. Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: results from TROG 02.02. J Clin Oncol 2010;28:2996–3001.
[13] Abrams RA, Winter KA, Regine WF, Safran H, Hoffman JP, Lustig R, et al. Failure to adhere to protocol specified radiation therapy guidelines was associated with decreased survival in RTOG 9704–a phase III trial of adjuvant chemotherapy and chemoradiotherapy for patients with resected adenocarcinoma of the pancreas. Int J Radiat Oncol Biol Phys 2012;82:809–16.
[14] Yuan L, Ge Y, Lee WR, Yin FF, Kirkpatrick JP, Wu QJ. Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans. Med Phys 2012;39:6868–78.
[15] Fogliata A, Belosi F, Clivio A, Navarria P, Nicolini G, Scorsetti M, et al. On the pre-clinical validation of a commercial model-based optimisation engine: application to volumetric modulated arc therapy for patients with lung or prostate cancer. Radiother Oncol 2014;113:385–91.
[16] Hussein M, South CP, Barry MA, Adams EJ, Jordan TJ, Stewart AJ, et al. Clinical validation and benchmarking of knowledge-based IMRT and VMAT treatment planning in pelvic anatomy. Radiother Oncol 2016;120:473–9.
[17] Xhaferllari I, Wong E, Bzdusek K, Lock M, Chen JZ. Automated IMRT planning with regional optimization using planning scripts. J Appl Clin Med Phys 2013;14:176–91.
[18] Hansen CR, Bertelsen A, Hazell I, Zukauskaite R, Gyldenkerne N, Johansen J, et al. Automatic treatment planning improves the clinical quality of head and neck cancer treatment plans. Clin Transl Radiat Oncol 2016;1:2–8.
[19] Vanderstraeten B, Goddeeris B, Vandecasteele K, van Eijkeren M, De Wagter C, Lievens Y. Automated instead of manual treatment planning? A plan comparison based on dose-volume statistics and clinical preference. Int J Radiat Oncol Biol Phys 2018;102:443–50.
[20] Purdie TG, Dinniwell RE, Fyles A, Sharpe MB. Automation and intensity modulated radiation therapy for individualized high-quality tangent breast treatment plans. Int J Radiat Oncol Biol Phys 2014;90:688–95.

[21] Mizuno N, Yamauchi R, Kawamori J, Itazawa T, Shimbo M, Nishimura K, et al. Evaluation of a new commercial automated planning software for tangential breast intensity-modulated radiation therapy. Radiol Phys Technol 2019;12:249–59.
[22] Zarepisheh M, Hong L, Zhou Y, Oh JH, Mechalakos JG, Hunt MA, et al. Automated intensity modulated treatment planning: the expedited constrained hierarchical optimization (ECHO) system. Med Phys 2019;46:2944–54.
[23] Fiandra C, Rossi L, Alparone A, Zara S, Vecchi C, Sardo A, et al. Automatic genetic planning for volumetric modulated arc therapy: a large multi-centre validation for prostate cancer. Radiother Oncol 2020;148:126–32.
[24] Craft D, Halabi T, Shih HA, Bortfeld T. An approach for practical multiobjective IMRT treatment planning. Int J Radiat Oncol Biol Phys 2007;69:1600–7.
[25] Craft DL, Hong TS, Shih HA, Bortfeld TR. Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy. Int J Radiat Oncol Biol Phys 2012;82:e83–90.
[26] Thieke C, Küfer K-H, Monz M, Scherrer A, Alonso F, Oelfke U, et al. A new concept for interactive radiotherapy planning with multicriteria optimization: first clinical evaluation. Radiother Oncol 2007;85:292–8.
[27] Kamran SC, Mueller BS, Paetzold P, Dunlap J, Niemierko A, Bortfeld T, et al. Multi-criteria optimization achieves superior normal tissue sparing in a planning study of intensity-modulated radiation therapy for RTOG 1308-eligible non-small cell lung cancer patients. Radiother Oncol 2016;118:515–20.
[28] Breedveld S, Storchi PRM, Voet PWJ, Heijmen BJM. iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. Med Phys 2012;39:951–63.
[29] Buergy D, Sharfo AWM, Heijmen BJM, Voet PWJ, Breedveld S, Wenz F, et al. Fully automated treatment planning of spinal metastases - a comparison to manual planning of Volumetric Modulated Arc Therapy for conventionally fractionated irradiation. Radiat Oncol 2017;12.
[30] Sharfo AW, Breedveld S, Voet PW, Heijkoop ST, Mens JM, Hoogeman MS, et al. Validation of fully automated VMAT plan generation for library-based plan-of-the-day cervical cancer radiotherapy. PLoS One. 2016;11:e0169202.
[31] Voet PWJ, Dirkx MLP, Breedveld S, Al-Mamgani A, Incrocci L, Heijmen BJM. Fully automated volumetric modulated arc therapy plan generation for prostate cancer patients. Int J Radiat Oncol Biol Phys 2014;88:1175–9.
[32] Voet PWJ, Dirkx MLP, Breedveld S, Fransen D, Levendag PC, Heijmen BJM. Toward fully automated multicriterial plan generation: a prospective clinical study. Int J Radiat Oncol Biol Phys 2013;85:866–72.
[33] Buschmann M, Sharfo AWM, Penninkhof J, Seppenwoolde Y, Goldner G, Georg D, et al. Automated volumetric modulated arc therapy planning for whole pelvic prostate radiotherapyAutomatisierte volumenmodulierte Arc-Therapieplanung für Ganzbecken-Prostatabestrahlung. Strahlenther Onkol 2018;194:333–42.
[34] Sharfo AWM, Dirkx MLP, Bijman RG, Schillemans W, Breedveld S, Aluwini S, et al. Late toxicity in the randomized multicenter HYPRO trial for prostate cancer analyzed with automated treatment planning. Radiother Oncol 2018;128:349–56.
[35] Sharfo AWM, Stieler F, Kupfer O, Heijmen BJM, Dirkx MLP, Breedveld S, et al. Automated VMAT planning for postoperative adjuvant treatment of advanced gastric cancer. Radiat Oncol 2018;13.
[36] Bijman R, Rossi L, Janssen T, de Ruiter P, Carbaat C, van Triest B, et al. First system for fully-automated multi-criterial treatment planning for a high-magnetic field MR-Linac applied to rectal cancer. Acta Oncol 2020;59:926–32.
[37] Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. Med Phys 1998;25:656–61.
[38] Ezzell GA, Burmeister JW, Dogan N, LoSasso TJ, Mechalakos JG, Mihailidis D, et al. IMRT commissioning: multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119. Med Phys 2009;36:5359–73.
[39] Aluwini S, Pos F, Schimmel E, Krol S, van der Toorn PP, de Jager H, et al. Hypofractionated versus conventionally fractionated radiotherapy for patients with prostate cancer (HYPRO): late toxicity results from a randomised, non-inferiority, phase 3 trial. Lancet Oncol 2016;17:464–74.
[40] Rossi L, Breedveld S, Heijmen BJM, Voet PWJ, Lanconelli N, Aluwini S. On the beam direction search space in computerized non-coplanar beam angle optimization for IMRT-prostate SBRT. Phys Med Biol 2012;57:5441–58.
[41] Hansen CR, Crijns W, Hussein M, Rossi L, Gallego P, Verbakel W, et al. Radiotherapy Treatment plannINg study Guidelines (RATING): a framework for setting up and reporting on scientific treatment planning studies. Radiother Oncol 2020;153:67–78.