Physics and Imaging in Radiation Oncology 24 (2022) 7–13

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

journal homepage: www.sciencedirect.com/journal/physics-and-imaging-in-radiation-oncology

Original Research Article

Benchmarking daily adaptation using fully automated radiotherapy treatment plan optimization for rectal cancer

Thyrza Z. Jagt, Tomas M. Janssen, Anja Betgen, Lisa Wiersema, Rick Verhage, Sanne Garritsen, Tineke Vijlbrief-Bosman, Peter de Ruiter, Peter Remeijer, Corrie A.M. Marijnen, Femke P. Peters, Jan-Jakob Sonke *

Department of Radiation Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands

ARTICLE INFO

Keywords:
Online plan adaptation
MRI guided radiotherapy
Automated treatment planning

ABSTRACT

Background/purpose: In daily plan adaptation the radiotherapy treatment plan is adjusted just prior to delivery. A simple approach is taking the planning objectives of the reference plan and directly applying these in re-optimization. Here we present a tested method to verify whether daily adaptation without tweaking of the objectives can maintain the plan quality throughout treatment.

Materials/methods: For fifteen rectal cancer patients, automated treatment planning was used to generate plans mimicking manual reference plans on the planning scans. For 74 fraction scans (4–5 per patient) an automated plan and a daily adapted plan were generated, where the latter re-optimizes the reference plan objectives without any tweaking. To evaluate the robustness of the daily adaptation, the adapted plans were compared to the autoplanning plans.

Results: Median differences between the autoplanning plans on the planning scans and the reference plans were between −1 and 0.2 Gy. The largest interquartile range (1 Gy) was seen for the Lumbar Skin D2%. For the daily scans the PTV D2% and D98% differences between autoplanning and adapted plans were within ±0.7 Gy, with mean differences within ±0.3 Gy. Positive differences indicate higher values were obtained using autoplanning. For the Bowelarea + Bladder and the Lumbar Skin the D2% and Dmean differences were all within ±2.6 Gy, with mean differences between −0.9 and 0.1 Gy.

Conclusion: Automated treatment planning can be used to benchmark daily adaptation techniques. The investigated adaptation workflow can robustly perform high quality adaptations without daily adjusting of the patient-specific planning objectives for rectal cancer radiotherapy.

1. Introduction

Daily online plan adaptation, where the plan is adjusted prior to delivering the treatment, is increasingly becoming available for radiotherapy [1–7]. In full plan adaptation, the tumor and organs at risk (OARs) are re-contoured and the plan is re-optimized to fit the anatomy of the day.

In online plan adaptation, a key question is how to effectively obtain an optimized treatment plan. A common approach is to take the planning objectives of the reference plan. Typically, these planning objectives are copied to the daily scan, where the onsite expert planner can manually tweak them to create an entirely new daily plan [1,2,4,6,7]. The plan quality can be validated by evaluating a set of predefined clinical criteria regarding dose and plan complexity. This approach requires experts to be present at the treatment machine and the iterative nature will prolong the adaptation time. A better approach would be to determine a set of planning objectives only once, and apply these to all treatment fractions. Optimizing the plan on the daily scans without modifying the objectives saves time and resources, leading to a more efficient daily workflow. In this workflow, however, it is uncertain whether manual tweaking could have resulted in a better plan. The question thus becomes how to verify whether this non-tweaking approach maintains the plan quality of the reference plan throughout the treatment. To our knowledge no studies have been done on the validation of such clinical adaptation techniques.

Fully automated treatment planning can generate high quality

* Corresponding author at: The Netherlands Cancer Institute, Department of Radiation Oncology, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.
E-mail address: j.sonke@nki.nl (J.-J. Sonke).

https://doi.org/10.1016/j.phro.2022.08.006
Received 31 March 2022; Received in revised form 11 August 2022; Accepted 11 August 2022
Available online 18 August 2022
2405-6316/© 2022 Published by Elsevier B.V. on behalf of European Society of Radiotherapy & Oncology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

treatment plans with consistent trade-offs for different patients [8–15]. With minimal user interaction, these methods are very suitable for treatment planning benchmark studies. By tuning a fully automated treatment planning technique to mimic the manually made reference plans, autoplanning can be applied to benchmark the daily plan adaptation technique.

In this paper we present a method to verify whether the treatment quality can be maintained throughout treatment without daily tweaking of the objective functions. As an example, we applied the method for plan adaptation for rectal cancer patients by simulating clinical treatments and benchmarking these treatments against fully optimized treatment plans.

2. Methods/Materials

2.1. Daily adaptation

The Unity MR-Linac has been equipped with an adapt-to-shape (ATS) workflow for daily full plan adaptation [16–19]. To run this workflow, a reference plan was generated on the planning scan. These reference plans were generated through manual tweaking of the objective functions of a site specific template, taking around 2–3 h each. For each fraction the ATS workflow applied the reference planning objectives to the corresponding region of interest on the daily redelineated new anatomy.

2.2. Patient data

Data of fifteen rectal cancer patients who received short course radiotherapy (5x5Gy; n = 5) or long course (chemo)radiotherapy (25x2Gy; n = 10) on the Unity MR-Linac between October 2018 and March 2021 were included in this study. These were the first fifteen patients from this time period for whom all required structures were delineated by experienced RTTs. For all patients one planning CT scan and 4–25 T2-weighted daily MRIs were acquired during the course of the treatment. For one short-course patient, only four MRIs were available as one of the fractions was delivered on a regular Linac. Per patient, five daily MRIs (consecutive for short-course patients, one per week for long-course patients) were selected for the analysis, yielding a total of 74 daily MRIs. Approval from the institutional review board was obtained.

2.3. Treatment volumes and dose prescription

For our study, 5x5Gy treatment plans were (re-)optimized for all planning scans by expert planners using the research version of the treatment planning system (TPS). Table 1 provides definitions of the delineated targets and OARs. An example is shown in Fig. 1. Dose was prescribed to the total PTV.

All plans were generated in the Monaco TPS (Elekta AB, Stockholm, Sweden, research version 5.59.11a).

2.4. Planning methods

All reference plans on the planning scans were generated through manual tweaking of a premade planning template (Tables S1 and S2 of the Supplementary Materials).

A reference plan was deemed acceptable when for all PTVs the V95% > 99 %, and for the total PTV D1% < 107 % and the Dmean was between 100 and 101 %. For the OARs, no hard criteria were set. Instead planning objectives on the BO+BL were applied to minimize the volume receiving over 22 Gy, and aim for a gradual dose fall off by pushing down the DVH curve while avoiding lateral hotspots. More details are shown in Table S1.

The ATS workflow takes the objectives used to obtain the reference plan on the planning CT and uses these to optimize a new plan on the daily MRI [16]. In this study adaptation was performed without any manual tweaking, so using a single optimization. To run the dose calculation the structures on the daily MR were assigned relative electron densities equal to the average values of the planning CT for the CTVs, Bones and BO+BL. The remaining area was assigned a relative electron density of 1.

Table 1
Targets and OARs were accurately delineated on all scans. The targets were delineated according to national guidelines adapted from Valentini et al. [20].

Targets	
CTV_{mps}	Clinical target volumes of the mesorectum and pre-sacral lymph node region.
CTV_{ln_L} and CTV_{ln_R} (left and right, respectively)	Clinical target volumes of the lateral lymph node regions (including internal iliac in all patients and on indication the obturator region).
PTV_{mps}	The CTV_{mps} expanded using an anisotropic margin of 8 mm in the anterior direction and 5 mm in all other directions.
PTV_{ln_L} and PTV_{ln_R} (left and right, respectively)	The CTV_{ln_L} and CTV_{ln_R} expanded using an isotropic margin of 5 mm.
PTV	$PTV_{mps} + PTV_{ln\_L} + PTV_{ln\_R}$, clipped at 6 mm from the External contour.
OARs	
BO+BL	The combined area of the bladder and bowel area.
Lumbar Skin	The 1 cm dorsal region behind the PTV [21], included to account for the close proximity of the PTV to the dorsal skin and the electron return effect caused by the magnetic field [22].

According to institutional protocol, an adapted plan was deemed clinically acceptable when the PTV D2% and D98% values were at most 0.5 Gy worse than the reference plan values. These daily criteria are slightly less strict than the reference plan criteria, as due to the daily replanning any underdosage is assumed to become random instead of systematic.

While no fully automated treatment planning is yet available for clinical use on the MR-Linac, the vendor Elekta AB has developed a feature to allow autoplanning in a research setting [8,21,23], called mCycle.

mCycle (based on Erasmus-iCycle [24,25]) allows for automated plan optimization based on a wishlist holding planning constraints and objectives. Using this wishlist mCycle runs a 2-phase lexicographic fluence map optimization, followed by segmentation (see the Supplementary Materials for a brief explanation).

A wishlist was designed and approved by a clinician such that the mCycle plans mimic the trade-offs that are made in current clinical practice. To check the quality of the wishlist, mCycle plans were generated on the reference images and compared to the reference plans.

To benchmark the ATS workflow an mCycle plan was generated for each daily MRI. From here on ATS will be referred to as “adaptation” and mCycle will be “autoplanning”.

2.5. Evaluation of the planning methods

All plans on the planning scans were evaluated in the Monaco TPS. To verify if our autoplanning technique could be used to benchmark our adaptation method, we first compared the reference plans to the autoplans on the planning scans.

For the daily adapted plans the clinical acceptability was checked using the PTV $D98\%_{daily} > D98\%_{reference} - 0.5$ Gy and PTV $D2\%_{daily} < D2\%_{reference} + 0.5$ Gy criteria. For all daily plans we evaluated the PTV D98% and D2% and for the OARs the D2% and Dmean values. We furthermore calculated the compensated adapted and autoplan values by subtracting the respective planning values from the daily obtained values.

We also investigated whether the daily adaptation quality depends on the magnitude of change in the PTV between the reference scan and the daily scans. To do this we determined the Mean Surface Distance

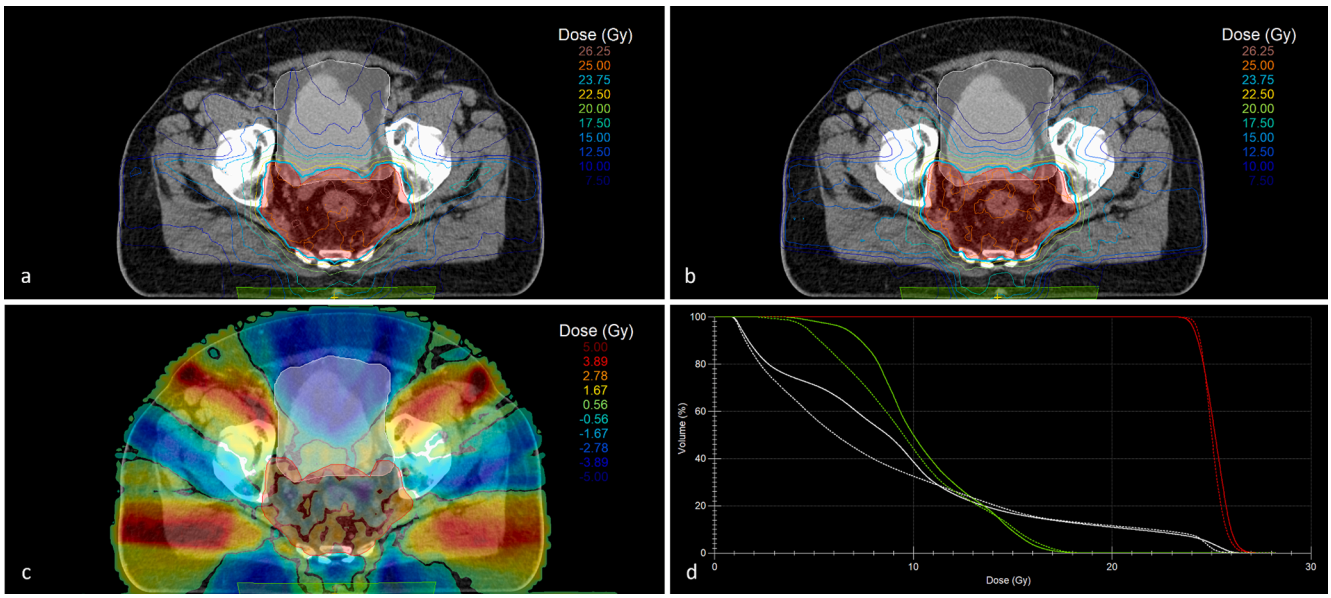

Fig. 1. A slice of the planning CT scan of one of the patients showing the reference dose distribution (a), the autoplan dose distribution (b), and the difference dose distribution (autoplan – reference) (c). The PTV is shown in red, the Lumbar Skin in green and the BO + BL in white. Subplot (d) shows the corresponding DVH curves. The solid line denotes the reference plan, the dashed line denotes the autoplan. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(MSD) for the PTV and the PTV D98% difference between the reference plan and the adapted plan for all fractions.

2.6. Statistical analysis

Differences between the adapted plans and autoplans were tested for statistical significance using the Wilcoxon signed-rank test for clustered data with significance level $\alpha = 0.05$ [26,27].

3. Results

3.1. Plan comparison on the planning scans

Table S3 of the Supplementary Materials shows the wishlist used to generate the autoplans.

All reference plans and autoplans generated for the planning scans were clinically acceptable. For the PTV D98%, the reference plans achieved a median value of 95.9 % (95.6 % – 97.1 %) and the autoplans achieved a median value of 96.7 % (95.6 % – 97.1 %) (see illustrative comparison in Fig. 1).

The difference between the autoplans and the reference plans for the PTV D98% was positive in median (0.1 Gy). All other median differences were between 0 and −1 Gy (see Fig. 2).

The largest interquartile range (IQR) of the dose differences was seen for the Lumbar Skin D2% (1 Gy). The smallest variation was seen for the PTV D98% (IQR 0.2 Gy).

Overall the median differences were close to zero and the small interquartile ranges showed that the autoplans could successfully mimic the reference plans. The autoplanning technique could hence be used to benchmark the adaptation workflow.

3.2. Plan comparison on the daily scans

Most PTV D98% differences between autoplanning and adaptation were positive (top left subplot of Fig. 3), indicating that autoplanning obtained slightly higher values than adaptation. The observed downward trend, shows that with an increasing average value the difference value decreased. For the PTV D2% the opposite was observed; most of

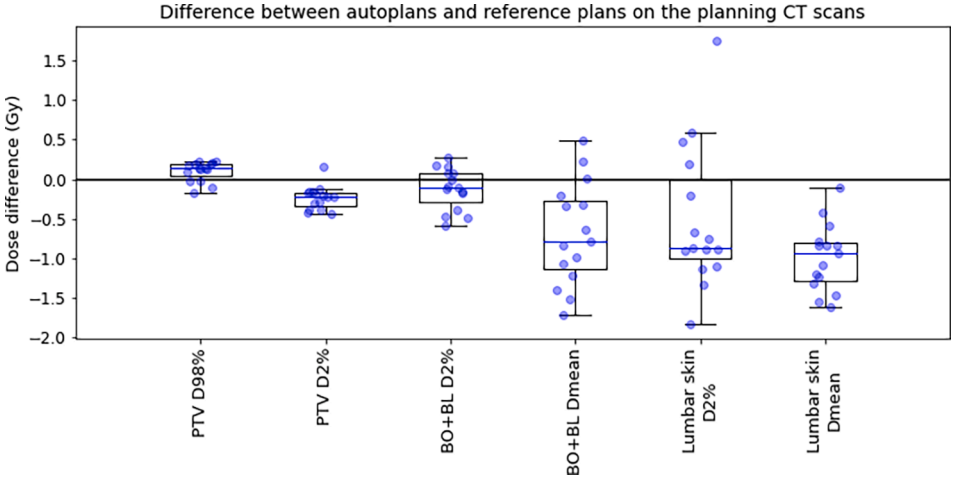

Fig. 2. A boxplot depicting the differences between the dosimetric values obtained on the planning CT scan using autoplanning, and the values obtained in the reference plan. Positive values indicate a higher value was obtained in the autoplan. The blue dots represent the individual plan values. Each box indicates the median and the 25th and 75th percentiles of the obtained differences. The vertical whiskers depict the remaining points up to 1.5 times the interquartile range. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

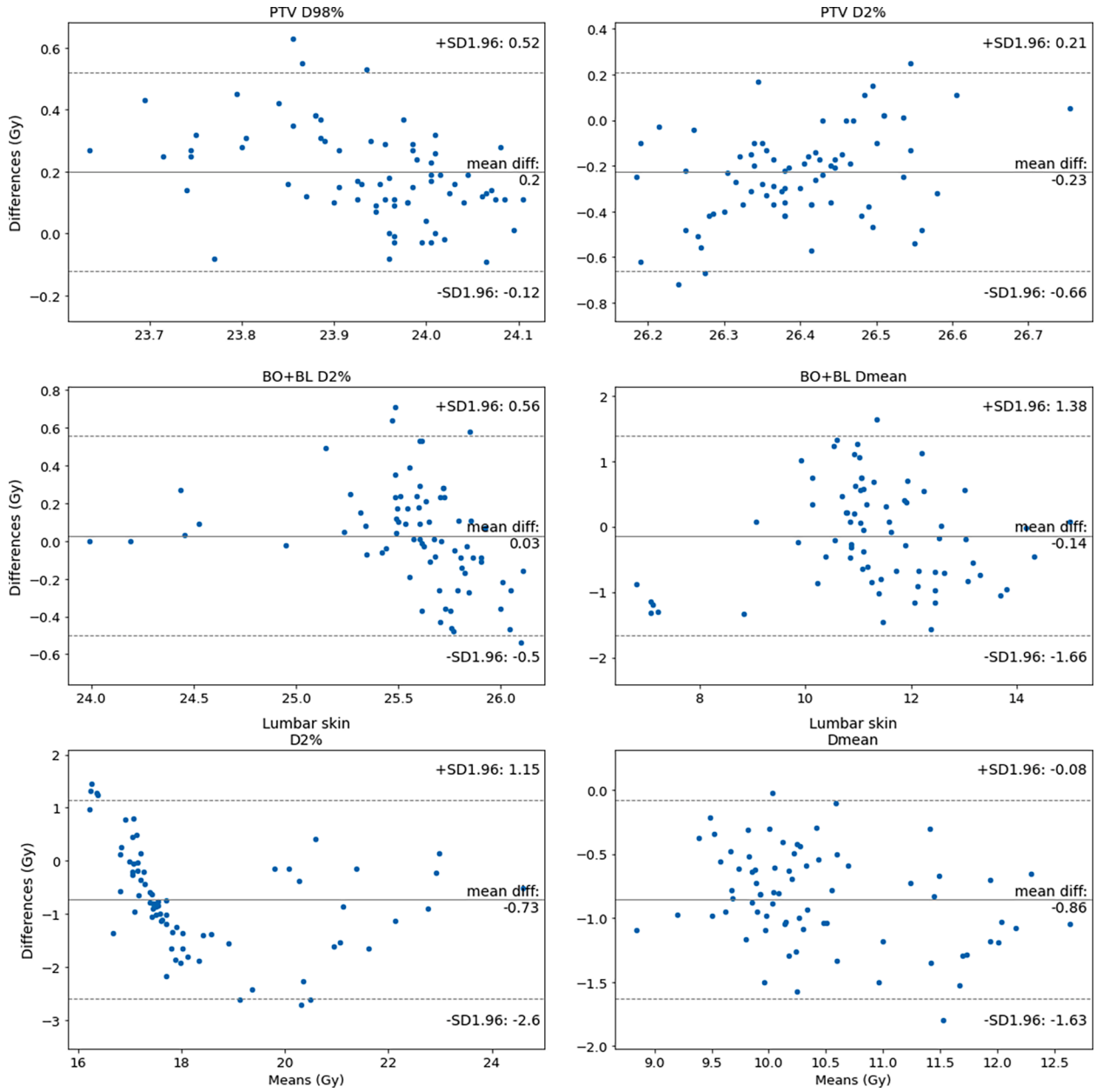

Fig. 3. Bland-Altman plots for all the dosimetric criteria showing the values obtained for all daily MRI scans using autoplanning and the adaptation workflow. Each dot represents a daily MRI scan: the x-value shows the average of the two plans, and the y-value the difference between the two plans. Shown differences are autoplanning – adaptation, i.e. positive differences indicate higher values were obtained using autoplanning. The limits of agreement (LoA) are plotted at ±1.96SD.

the differences were negative and a slight upward trend was seen (top right subplot).

The clinical criterion PTV $D98\%_{daily} > D98\%_{reference} - 0.5$ Gy was met for 69/74 of the adapted plans. Five adapted plans got values between 0.5 and 0.9 Gy lower than reference. For all 74 scans the daily PTV D2% was less than 0.5 Gy worse than the reference value.

A PTV D98%>95 % was achieved for 53/74 and 73/74 plans for adaptation and autoplanning, respectively. For the adapted plans the lowest achieved PTV D98% was 93.4 %, for the autoplans this was 94.9 %, for a different patient. The PTV D2% values were below 107 % for 72/74 adapted and 72/74 autoplans. Highest values were 107.3 % and 107.6 % for adaptation and autoplanning respectively.

Fig. 3 shows that for the OARs the differences between the adapted and autoplans all stayed within 2.6 Gy, with mean differences between −0.9 and 0.1. For the BO+BL D2% the LoA interval indicates that differences mostly stayed within 0.6 Gy, while more variation was seen for the BO+BL Dmean. The largest variation was seen for the Lumbar Skin D2%, where the LoA indicated a 95 % difference interval between −2.6 and 1.2 Gy. For this criteria some correlation was seen for the plans with low doses, showing that autoplanning did not push quite as hard to reduce doses. For the Lumbar Skin Dmean all differences were negative, meaning that the adaptation obtained higher values for all fractions of all patients.

Overall the differences in OAR doses between adapted and autoplans were similar to those observed on the planning scans, suggesting that they primarily originated from differences between the planning techniques rather than the quality of the patient-specific selected planning objectives. The differences between the compensated adaptation and autoplanning values were only statistically significant for the BO+BL D2% ($p = 0.04$) and BO+BL Dmean ($p = 0.01$).

Fig. 4 depicts the Mean Surface Distance values plotted against the differences in PTV D98% between the reference plan and adapted plan

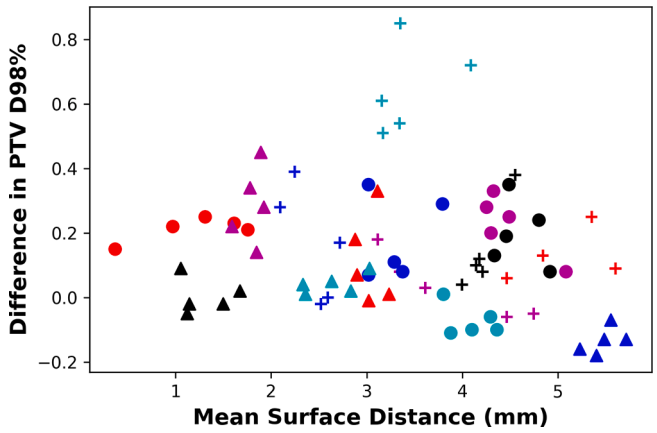

Fig. 4. The mean surface distance between the reference scan PTV and the daily PTV is plotted against the difference in PTV D98% value between the reference plan and the adapted plan for all fractions of all patients. A positive D98% difference indicates a higher PTV D98% value was obtained in the reference plan. Different patients are indicated by different color and symbol combinations.

for all fractions.

3.3. Timing results

Differences between plan properties and timing results on both the planning scans and the daily scans were statistically significant for all but the segmentation time (Table 2).

4. Discussion

In this study we have described a general method to verify whether plan quality can be maintained throughout treatment without daily manual tweaking of the planning objectives. By tuning an automated planning technique to mimic clinical plan quality, this autoplanning tool can be used to benchmark daily plan adaptation.

As an example we have benchmarked our clinical ATS workflow for rectal cancer treatment on the Unity MR-Linac against the research autoplanning technique mCycle. We found that autoplanning (mCycle) can be used to mimic the clinical plan quality of the reference plans, and that directly applying the reference planning objectives in the ATS adaptation workflow generally resulted in clinically acceptable daily plans of comparable quality to fully optimized autoplans.

To apply the proposed method, a set of patients with repeat scans and daily delineations is required. The methodology then consists of first ensuring that the autoplanning technique generates plans with similar trade-offs as the clinical plans. When verified, the autoplanning technique can be applied to the daily scans. The resulting plans can then be compared to the adapted plans to evaluate the adaptation quality. If the autoplanning technique and the clinical planning system use different optimization techniques, achieving exactly the same trade-offs will be challenging. Small differences between the automated and manual plans will then remain, yet the benchmarking technique can still be used to spot outliers. In this study such differences were mostly visible for the OARs, for which the autoplanning technique achieved a bit more sparing than was seen in the clinical reference plans (Fig. 1). Similar differences, i.e. slightly higher target coverages and lower OAR doses, were also observed in mCycle validation studies [8,21,23].

In our practical example the adaptation workflow directly resulted in clinically acceptable plans for 69/74 scans. The other 5 scans were all from the same patient. Clinically these plans were fixed through tweaking of the planning objectives (first fraction), and using a new reference plan made on the manually delineated first fraction MRI (other fractions). As the new reference plan proved adequate for the remaining fractions, the lower target coverage observed in this study is likely explained by a large overall difference between the planning anatomy and the first fraction anatomy. This study thus shows that most patients can be treated using this adaptation workflow. Using a threshold in evaluating the daily adapted plan can ensure an effective workflow; recognizing outliers to only adjust the daily planning objectives when necessary. While we used data of only 15 patients, we do not expect the conclusions would change if more patients would be included.

The negative mean differences for the OARs (Fig. 3) show that on the daily scans, similar to on the planning scans, autoplanning generally resulted in slightly lower values than adaptation. After compensating for the differences between the two planning methods, the differences between the daily plans were only statistically significant for the BO+BL D2% and Dmean. This could be due to the trade-off between the target coverage and BO+BL dose. While for the planning scans all plans achieved an acceptable coverage, 21/74 adapted plans achieved a D98%< 95 % vs 1/74 autoplans. The higher target coverage in autoplanning could also result in higher doses to the BO+BL.

Nijkamp et al. [28] (Table 2) showed that for 5x5Gy conventional treatments 7/15 mm margins are minimally required to assure sufficient target coverage in at least 90 % of the patients. In this study all plans were generated using 5/8 mm margins, only accounting for intrafraction motion and delineation uncertainties. As our used margins are a lot tighter than those required for conventional treatments, daily adaptation is thus required.

No correlation could be seen between the MSD for the PTV and the differences in PTV D98% in Fig. 4, indicating that the quality of the adaptation is independent of the amount of anatomical change in the

Table 2
Median number of segments and monitor units obtained in the adapted and autoplans and the timing results of the different optimization steps in the reference and autoplanning optimizations for the planning scans and the adaptation and autoplanning optimizations for the daily MRI scans. Segmentation includes the segment shape optimization (SSO) and the segment weight optimization (SWO). The full optimization includes everything from start to finish, hence including some overhead steps on top of the fluence map optimization and segmentation.

	Planning Scans		Daily Scans	
	Reference	Autoplan	Adapted	Autoplan
# segments Median*(min–max)*	69 (67–80)	66 (47–78)	68 (49–79)	67 (45–79)
# MU Median*(min–max)*	1535 (1285–1747)	1808 (1578–2217)	1757 (1181–2039)	1958 (1485–2362)
Fluence map optimization (minutes) Median*(min–max)*	0.1 (0.1–0.3)	2.1 (1.4–3.4)	0.2 (0.1–0.3)	2.5 (1.5–6.0)
Segmentation (minutes) Median*(min–max)*	5.2 (4.4–6.5)	5.3 (3.8–6.4)	4.4 (3.4–5.2)	4.4 (3.3–5.3)
Full optimization (minutes) Median*(min–max)*	5.5 (4.6–6.8)	7.5 (5.3–9.3)	4.7 (3.6–5.4)	7.1 (5.2–10.4)

PTV. Datapoints were also clustered per patient, indicating more inter-patient than intrapatient variation.

In this study we investigated the robustness of the adapt-to-shape workflow on the MR-Linac without daily tweaking of the planning objectives for rectal cancer patients. In a similar study, Intven et al. [1] looked into 5x5Gy ATS treatments for rectal cancer patients using 5 beams. Their result of achieving acceptable target coverage match our own results, our study furthermore investigated whether improvements were possible for both targets and OARs. Similarly Winkel et al. [29] looked into daily adaptation for stereotactic treatments of lymph node oligometastases. While they show that from the different adaptation techniques the ATS from fluence yields the best results, no evaluation of the quality compared to manual planning was performed.

The main reason to avoid manual tweaking is the reduction of required personnel resources and adaptation times. Table 2 shows that in median the adaptation optimization took about 4.5 min, autoplanning about 7. It should be noted that these calculations have been performed on research hardware holding a single NVIDIA GeForce RTX 2080 TI GPU card, while our current clinical hardware contains two NVIDIA Quadro GP100 GPU cards and hence will be faster. Note that if the investigated autoplanning technique would be clinically available, the adaptation technique could thus be replaced without much time loss. This study however shows that the daily dosimetric gain would also be limited for this patient group. In generating the reference plans the use of autoplanning could nevertheless lead to a substantial time reduction, as the manual tweaking will no longer be necessary. It should be noted though that the autoplans in this study used more Monitor Units than the reference/adapted plans (Table 2), which will result in an increase in delivery time. A median autoplan (67 segments, 1958 MU, 8 min 15 s) will take approximately 27 s longer than a median adapted plan (68 segments, 1757 MU, 7 min 48 s). As even the worst case scenario of 79 segments and 2362 MU could be delivered within 10 min, all of these plans would be deemed clinically acceptable in our institute.

In this study all reference plans were made by experienced planners. While we realize that the adaptation results depend on the set of chosen objectives for the reference plan, predicting what set of objectives would be most robust for daily adaptation remains difficult. The robustness of the set of planning objectives may furthermore be TPS dependent. Monaco might have an advantage over other systems, as it uses constraint-optimization and automatically applies any shrink-margins that were applied in the reference plan to the daily structures. Another limitation is that as we considered rectal cancer patients, we only included two OARs. One patient experienced large anatomical variations between the planning scan and the first fraction scan, causing the adaptation workflow to yield inadequate target coverage. For other treatment sites with more OAR objectives this could possibly happen more often.

In conclusion, this study has provided a general method to verify whether a set of planning objectives can be used throughout the entire treatment of a patient, by benchmarking the adaptation technique against fully automated plans. As an example we have applied the method to show that a relatively simple adaptation workflow can robustly perform high quality adaptations without daily adjusting of the patient-specific planning objectives.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Netherlands Cancer Institute is a member of the Elekta AB (Stockholm, Sweden) MR-Linac consortium. This work uses TPS Monaco Research version 5.59.11a, which is made available to the Netherlands Cancer Institute as part of a research collaboration.

Acknowledgements

The authors want to thank Peter Voet and Randy Larson from Elekta for their technical support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phro.2022.08.006.

References

[1] Intven MPW, de Mol van Otterloo SR, Mook S, Doornaert PAH, de Groot-van Breugel EN, Sikkes GG, et al. Online adaptive MR-guided radiotherapy for rectal cancer; feasibility of the workflow on a 1.5T MR-linac: clinical implementation and initial experience. Radiother Oncol 2021;154:172–8. https://doi.org/10.1016/j.radonc.2020.09.024.

[2] Finazzi T, Palacios MA, Spoelstra FOB, Haasbeek CJA, Bruynzeel AME, Slotman BJ, et al. Role of on-table plan adaptation in MR-guided ablative radiation therapy for central lung tumors. Int J Radiat Oncol Biol Phys 2019;104:933–41. https://doi.org/10.1016/j.ijrobp.2019.03.035.

[3] Liu X, Liang Y, Zhu J, Yu G, Yu Y, Cao Q, et al. A fast online replanning algorithm based on intensity field projection for adaptive radiotherapy. Front Oncol 2020;10:1–10. https://doi.org/10.3389/fonc.2020.00287.

[4] Henke L, Kashani R, Robinson C, Curcuru A, DeWees T, Bradley J, et al. Phase I trial of stereotactic MR-guided online adaptive radiation therapy (SMART) for the treatment of oligometastatic or unresectable primary malignancies of the abdomen. Radiother Oncol 2018;126:519–26. https://doi.org/10.1016/j.radonc.2017.11.032.

[5] Bernatowicz K, Geets X, Barragan A, Janssens G, Souris K, Sterpin E. Feasibility of online IMPT adaptation using fast, automatic and robust dose restoration. Phys Med Biol 2018;63:085018. https://doi.org/10.1088/1361-6560/aaba8c.

[6] de Jong R, Visser J, van Wieringen N, Wiersma J, Geijsen D, Bel A. Feasibility of Conebeam CT-based online adaptive radiotherapy for neoadjuvant treatment of rectal cancer. Radiat Oncol 2021;16:1–11. https://doi.org/10.1186/s13014-021-01866-7.

[7] Van Timmeren JE, Chamberlain M, Krayenbuehl J, Wilke L, Ehrbar S, Bogowicz M, et al. Treatment plan quality during online adaptive re-planning. Radiat Oncol 2020;15:1–11. https://doi.org/10.1186/s13014-020-01641-0.

[8] Bijman R, Sharfo AW, Rossi L, Breedveld S, Heijmen B. Pre-clinical validation of a novel system for fully-automated treatment planning. Radiother Oncol 2021;158:253–61. https://doi.org/10.1016/j.radonc.2021.03.003.

[9] Schipaanboord BWK, Giżyńska MK, Rossi L, de Vries KC, Heijmen BJM, Breedveld S. Fully automated treatment planning for MLC-based robotic radiotherapy. Med Phys 2021;48:4139–47. https://doi.org/10.1002/mp.14993.

[10] Babier A, Mahmood R, McNiven AL, Diamant A, Chan TCY. Knowledge-based automated planning with three-dimensional generative adversarial networks. Med Phys 2020;47:297–306. https://doi.org/10.1002/mp.13896.

[11] Wang H, Wang R, Liu J, Zhang J, Yao K, Yue H, et al. Tree-based exploration of the optimization objectives for automatic cervical cancer IMRT treatment planning. Br J Radiol 2021;94:20210214. https://doi.org/10.1259/bjr.20210214.

[12] Zhang X, Li X, Quan EM, Pan X, Li Y. A methodology for automatic intensity-modulated radiation treatment planning for lung cancer. Phys Med Biol 2011;56:3873–93. https://doi.org/10.1088/0031-9155/56/13/009.

[13] 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. https://doi.org/10.1016/j.ijrobp.2013.12.046.

[14] Kusters JMAM, Bzdusek K, Kumar P, van Kollenburg PGM, Kunze-Busch MC, Wendling M, et al. Automated IMRT planning in Pinnacle: A study in head-and-neck cancer. Strahlenther Onkol 2017;193:1031–8. https://doi.org/10.1007/s00066-017-1187-9.

[15] Kusters JMAM, Bzdusek K, Kumar P, van Kollenburg PGM, Kunze-Busch MC, Wendling M, et al. Correction to: Automated IMRT planning in Pinnacle—A study in head-and-neck cancer (Strahlenther Onkol, (2017), 10.1007/s00066-017-1187-9). Strahlenther Onkol 2017;193:1077–8. https://doi.org/10.1007/s00066-017-1230-x.

[16] Winkel D, Bol GH, Kroon PS, van Asselen B, Hackett SS, Werensteijn-Honingh AM, et al. Adaptive radiotherapy: The Elekta Unity MR-linac concept. Clin Transl Radiat Oncol 2019;18:54–9. https://doi.org/10.1016/j.ctro.2019.04.001.

[17] Kontaxis C, Bol GH, Lagendijk JJW, Raaymakers BW. A new methodology for inter- and intrafraction plan adaptation for the MR-linac. Phys Med Biol 2015;60:7485–97. https://doi.org/10.1088/0031-9155/60/19/7485.

[18] Raaymakers BW, Lagendijk JJW, Overweg J, Kok JGM, Raaijmakers AJE, Kerkhof EM, et al. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: Proof of concept. Phys Med Biol 2009;54:N229–37. https://doi.org/10.1088/0031-9155/54/12/N01.

[19] Lagendijk JJW, Raaymakers BW, van Vulpen M. The magnetic resonance imaging-linac system. Semin Radiat Oncol 2014;24:207–9. https://doi.org/10.1016/j.semradonc.2014.02.009.

[20] Valentini V, Gambacorta MA, Barbaro B, Chiloiro G, Coco C, Das P, et al. International consensus guidelines on Clinical Target Volume delineation in rectal

cancer. Radiother Oncol 2016;120:195–201. https://doi.org/10.1016/j.radonc.2016.07.017.

[21] 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. https://doi.org/10.1080/0284186X.2020.1766697.

[22] Raaijmakers AJE, Raaymakers BW, Lagendijk JJW. Integrating a MRI scanner with a 6 MV radiotherapy accelerator: Dose increase at tissue-air interfaces in a lateral magnetic field due to returning electrons. Phys Med Biol 2005;50:1363–76. https://doi.org/10.1088/0031-9155/50/7/002.

[23] Naccarato S, Rigo M, Pellegrini R, Voet P, Akhiat H, Gurrera D, et al. Automated planning for prostate stereotactic body radiation therapy on the 1.5 T MR-Linac. Adv Radiat Oncol 2022;7:1–12. https://doi.org/10.1016/j.adro.2021.100865.

[24] 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. https://doi.org/10.1118/1.3676689.

[25] 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. https://doi.org/10.1016/j.ijrobp.2012.04.015.

[26] Jiang Y, He X, Lee MLT, Rosner B, Yan J. Wilcoxon rank-based tests for clustered data with r package clusrank. J Stat Softw 2020;96:1–26. https://doi.org/10.18637/jss.v096.i06.

[27] Rosner B, Glynn RJ, Lee MLT. The Wilcoxon signed rank test for paired comparisons of clustered data. Biometrics 2006;62:185–92. https://doi.org/10.1111/j.1541-0420.2005.00389.x.

[28] Nijkamp J, Swellengrebel M, Hollmann B, De Jong R, Marijnen C, Van Vliet-Vroegindeweij C, et al. Repeat CT assessed CTV variation and PTV margins for short- and long-course pre-operative RT of rectal cancer. Radiother Oncol 2012;102:399–405. https://doi.org/10.1016/j.radonc.2011.11.011.

[29] Winkel D, Bol GH, Werensteijn-Honingh AM, Kiekebosch IH, van Asselen B, Intven MPW, et al. Evaluation of plan adaptation strategies for stereotactic radiotherapy of lymph node oligometastases using online magnetic resonance image guidance. Phys Imaging Radiat Oncol 2019;9:58–64. https://doi.org/10.1016/j.phro.2019.02.003.