



Original Research Article

Automatic treatment planning improves the clinical quality of head and neck cancer treatment plans



Christian Rønn Hansen^{a,*}, Anders Bertelsen^a, Irene Hazell^a, Ruta Zukauskaitė^{b,c}, Niels Gyldenkerne^c, Jørgen Johansen^{b,c}, Jesper G. Eriksen^{b,c}, Carsten Brink^{a,b}

^aLaboratory of Radiation Physics, Odense University Hospital, Odense, Denmark

^bInstitute of Clinical Research, University of Southern Denmark, Odense, Denmark

^cDepartment of Oncology, Odense University Hospital, Odense, Denmark

ARTICLE INFO

Article history:

Available online 19 September 2016

Keywords:

Automatic
Treatment planning
Head and neck
VMAT
Pinnacle

ABSTRACT

Background: Treatment plans for head and neck (H&N) cancer are highly complex due to multiple dose prescription levels and numerous organs at risk (OARs) close to the target. The plan quality is inter-planner dependent since it is dependent on the skills and experience of the dosimetrist. This study presents a blinded prospective clinical comparison of automatic (AU) and manually (MA) generated H&N VMAT plans made for clinical use.

Methods: MA and AU plans were generated for 30 consecutive patients in Pinnacle³ using the IMRT optimisation module and the new Autoplan module, respectively. The plan quality was blindly compared by three senior oncologists and the best plan was selected for treatment of the patient. Planning time was measured as the active operator time used. The plan quality was analysed with DVH metrics and the dose delivery accuracy validated on the ArcCheck phantom.

Results: For twenty-nine out of the thirty patients the AU plan was chosen for treatment. Target doses were more homogenous with the AU plans and the OAR doses were significantly reduced, between 0.5 and 6.5 Gy. The average operator time spent on creating a manual plan was 64 min which was halved by Autoplan. The AU plans were more modulated as illustrated by an increase in MUs, which might cause the slightly lower pass rate of 97.7% in the ArcCheck measurements.

Conclusions: Target doses were similar between MA and AU plan, while AU plans spared all OAR considerably better than the MA plans.

© 2016 The Authors. Published by Elsevier Ireland Ltd on behalf of European Society for Radiotherapy and Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Radiotherapy of head and neck (H&N) cancer is challenging due to complex shaped targets situated close to numerous radiosensitive organs at risk (OARs). Thus, a high quality H&N plan is conform, has homogeneous target doses according to prescription and as low doses to OAR's as possible with a proper balance between target coverage and OAR doses. The plan quality has been shown to directly affect treatment outcome in clinical trials, which in itself has created a new era of quality assurance [1,2].

Obtaining high quality plans is a demanding task. A high quality treatment plan relies on the skills and experience of the dosimetrist, which can vary greatly. Even for the most experienced

dosimetrist, time pressure to deliver a clinically applicable dose plan may result in treatment plans of inferior quality than desired. With several OARs in play in the head and neck region, there may be a tendency to focus more on specific OARs and thereby disregarding the importance of other risk organs during the optimisation process. This results in plans with acceptable target coverage and adequate sparing of e.g. the spinal cord and parotids, but for which e.g. the oral cavity and constrictor muscles are irradiated to higher levels than necessary [3]. This is not necessarily due to lack of skills or experience of the dosimetrist, but merely due to the ability to focus on only a limited number of objectives at a time.

The user-dependant variation of plan quality can be improved by applying specific guidelines which define the minimum standards for dose targets and OARs [4]. Such guidelines can to some extent reduce the overall variation of the dosimetrist-dependant plan quality, but they do not guide the dosimetrist towards the

* Corresponding author at: Odense University Hospital, Sdr. Boulevard 29, 5000 Odense, Denmark.

E-mail address: Christian.Roenn@rsyd.dk (C.R. Hansen).

optimal plan for the specific patient – more towards as guideline acceptable level.

Recently, new techniques have been applied for head and neck treatments such as Volume modulated arc therapy (VMAT) [5], Tomotherapy [6] and proton therapy [7]. These technical developments have increased the degrees of freedom compared to intensity modulated radiotherapy (IMRT) and thereby made it possible to improve the plan quality. Various approaches have been made for an automatic search for the optimal patient plan such as atlas-based planning (RapidPlan [8]), ideal dose distribution estimation (PlanIQ [9] and Erasmus-iCycle [10]), and template-based optimisation as used in AutoPlanning (AP) provided by Pinnacle³ [11,12].

In a previous study, we retrospectively evaluated a research version of AutoPlan in seven field step and shoot IMRT (ssIMRT) plans for H&N treatment. The results were promising based on DVH analysis as well as within a blinded clinical plan evaluation [12]. Since the study was made as a retrospective study, the plans delivery were not dosimetrically validated, and the plans were ssIMRT and not VMAT which is the current clinical standard in our department.

This prospective study presents a blinded comparison study of automatic and manually generated H&N VMAT plans made for clinical use utilizing the clinically released version of AutoPlan. Finally, the dosimetric accuracy and deliverability of the generated plans were validated.

2. Material and methods

All patients ($n = 30$) referred to curative H&N radiotherapy in August and September 2015 were included prospectively. The demographics (Table 1) are representative for the general entry of head and neck patients in the cancer centre, though with slightly more unknown primary tumours and less oral cavity tumours. No patients were censored from the study.

All patients followed the standard process of radiotherapy planning, i.e.: immobilisation, CT simulation (slice thickness of 3 mm and in plane voxel size of 1 mm × 1 mm), contouring by a radiologist and oncologist before the dosimetrist created a MA VMAT plan and an AU VMAT plan. Treatment planning was performed in accordance with the Danish Head and Neck Cancer Group's guidelines (DAHANCA – Version 2013 ver. 2.0 [13]) and each dose plan included up to three dose levels of 66 or 68 Gy (PTV1), 60 Gy (PTV2) and 50 Gy (PTV3) in 33 or 34 fractions. All plans used the simultaneous integrated boost technique and a full 360 degree VMAT arc for bilateral targets and a 200–220 degree VMAT arc for unilateral targets.

The MA plans were optimised according to standard clinical practice, with a script template for optimisation objectives. Manual adjustment of the template predefined objective values were performed a minimum of 3 times guided by the patient-specific possibilities as determined by the dosimetrist.

The AU plans were created by the AutoPlan software available in Pinnacle³ version 9.10, using settings shown in Table 2. The spinal cord and brain stem had higher priority than the target coverage (non-compromise); and the remaining OARs were automatically prioritised according to the proportion of overlap with the target (high if <25% overlap, medium if 25–50% overlap and low if >50% overlap). After AU optimisation, a minor manual fine-tuning of the plans was performed for all plans. No direct comparisons between the two plans were performed before both plans were finalised, however the bias of the same dosimetrist making both plans was minimised by one half of patients having the MA plans created first and for the other half of the patients having the AU plan was created first. Dose calculation was performed with

Table 1
Patient demographics.

| | Number | % |
|------------------------------------|--------|----|
| <i>Sex</i> | | |
| Male | 23 | 77 |
| Female | 7 | 23 |
| <i>WHO performance status</i> | | |
| 1 | 18 | 60 |
| 2 | 11 | 37 |
| 3 | 1 | 3 |
| <i>Site</i> | | |
| Oral cavity | 3 | 10 |
| Pharynx | 15 | 50 |
| Larynx | 4 | 13 |
| Salivary glands | 2 | 7 |
| Thyroid gland | 2 | 7 |
| Unknown primary tumour | 4 | 13 |
| <i>Stage</i> | | |
| II | 4 | 13 |
| III | 4 | 13 |
| IV | 22 | 73 |
| <i>Histology</i> | | |
| Squamous cell | 27 | 90 |
| Adenocarcinoma | 1 | 3 |
| Papillary carcinoma | 1 | 3 |
| Invasive ductal carcinoma | 1 | 3 |
| <i>HPV status (p16 expression)</i> | | |
| Positive | 12 | 40 |
| Negative | 12 | 40 |
| Irrelevant | 6 | 20 |
| <i>RT indication</i> | | |
| Definitive | 19 | 63 |
| Adjuvant | 11 | 37 |
| <i>Radiosensitiser*</i> | | |
| Yes | 19 | 63 |
| No | 11 | 37 |
| <i>Systemic treatment#</i> | | |
| Yes | 11 | 37 |
| No | 19 | 63 |
| <i>Target area</i> | | |
| Unilateral | 6 | 20 |
| Bilateral | 24 | 80 |

* As radiosensitiser, 1.2 g/m² Nimorazol was administered before every fraction.

Cisplatin 40 mg/m² weekly was used as systemic treatment.

Table 2
AutoPlan setup template.

| Organs at risk | Objective | Dose level | Compromise |
|---------------------|-----------|------------|------------|
| Brain stem | Max dose | 54 Gy | No |
| Brain stem PRV | Max dose | 60 Gy | No |
| Spinal cord | Max dose | 45 Gy | No |
| Spinal cord PRV | Max dose | 50 Gy | No |
| Parotid gland | Mean dose | 20 Gy | Yes |
| Submandibular gland | Mean dose | 35 Gy | Yes |
| Mandible | Max dose | 107% | Yes |
| Lips | Mean dose | 20 Gy | Yes |
| Thyroid gland | Mean dose | 40 Gy | Yes |
| Oral cavity | Mean dose | 30 Gy | Yes |
| Larynx | Mean dose | 44 Gy | Yes |
| Constrictor muscles | Mean dose | 45 Gy | Yes |
| Whole brain | Max dose | 60 Gy | Yes |

AutoPlan setup for all 30 patients. All values are adopted directly from the DAHANCA 2013 guidelines. The prioritisation was set automatically for all objectives as described in the text.

the Pinnacle³ collapsed cone algorithm with a dose grid resolution of 3 mm and a control point spacing of 2 degrees.

The planning techniques were blinded before MA and AU plans were presented at the daily clinical radiotherapy conference. Clinical evaluation was performed based on isodoses shown on

axial images, DVHs, and a protocol compliance scorecard derived from the DAHANCA guidelines. Plan selection was performed by three senior H&N oncologists and overviewed by oncologists of other treatment sites, medical physicists and RTT's.

To supplement the clinical evaluation, the operator time for the dosimetrist was recorded, and delivery accuracy of the plans was validated on an ArcCheck phantom using the pass rate of a gamma evaluation (3% of max measured dose, 3 mm).

Quantitative dosimetric evaluation of the treatment plans was performed on Dose Volume Histograms (DVHs) extracted from the planning system. The average DVH was calculated for each type of treatment plan as the average of the patient-specific DVH values at each dose level. The DVH analysis was performed for all target volumes as well as for the parotid glands, submandibular glands, the mandible, oral cavity, lips, larynx, thyroid, brain stem, and spinal cord. To evaluate radiation dose to the remaining healthy tissue, a DVH evaluation of all healthy tissues was performed. In contrast to the previous DVH evaluations, this evaluation was performed in absolute volume to compensate for a difference of the CT scanned volume of each patient (relative values would depend on the scanned volume).

The mean OAR doses and target conformity index ($CI = V_{95\%}/V_{PTV}$) and homogeneity index ($HI = D_{2\%} - D_{98\%}/D_{prescription}$) were calculated.

2.1. Statistics

Differences were tested using Wilcoxon matched-pair signed-rank with a significance level of 5%. To indicate dose regions in the DVHs for which statistically significant differences exist, a Wilcoxon matched-pair sign rank probability curve was calculated as in Bertelsen et al. [5].

3. Results

In 29 of the 30 plans, the AU plan was chosen for clinical application ($p < 0.001$).

All plans adhered to the critical objectives for targets and critical OAR (spinal cord and brain stem).

In terms of target coverage, the AU plans had a higher mean dose to all three PTVs. However, the D2 was lower for the AU plans. The HIs showed that each dose plateau was more homogenous for the AU plans compared to MA for each individual area of dose prescription. CI of the AU plans was significantly better for PTV3 (Table 3). According to ICRU83 [14], the PTV3 includes PTV2 causing the long tail towards high doses – likewise, PTV2 contains PTV1, however analysing PTV2only (PTV2–PTV1) and PTV3only (PTV3–PTV2), these overdose-tails were significantly reduced by AU compared to MA.

Mean OAR doses were significantly reduced in the AU plans for all risk organs. The mean reduction ranged from 0.5 Gy for the entire patient (body) to 6.5 Gy for the contralateral submandibular gland (Table 4). Differences in mean DVH showed significant AU superiority in the dose range 10–45 Gy for all organs. Average DVH are shown in Fig. 1 for six of the OAR and in Fig. 2 for four ring structures surrounding the targets. During the clinical plan evaluation, the general impression was that the AU plan was visually more conform, spared OAR better and had steeper dose fall away from the targets. The irradiation of OARs was in particular reduced for the submandibular glands. In Fig. 3 a screen dump of a representative patient is shown. Here the sparing of the right submandibular gland, the extended oral cavity and the spinal cord is mainly seen for the dose spillage iso-dose curves of 40 and 30 Gy.

The only manual plan selected for clinical use was a non-typical H&N case for a thyroid cancer involving level VII lymph nodes. That

Table 3
Target doses.

| | Autoplan | | Manual | | p |
|---------------------------|----------|------|--------|------|--------|
| | Mean | STD | Mean | STD | |
| <i>CTV1</i> | | | | | |
| Mean | 100.8 | 0.3 | 100.2 | 0.4 | <0.001 |
| D2 | 103.2 | 0.7 | 103.2 | 1.4 | <0.001 |
| D98 | 97.5 | 1.9 | 96.7 | 1.9 | <0.001 |
| <i>PTV1 (66 or 68 Gy)</i> | | | | | |
| Mean | 100.3 | 0.3 | 99.7 | 0.5 | <0.001 |
| D2 | 103.3 | 0.6 | 103.5 | 1.2 | <0.001 |
| D98 | 96.1 | 0.9 | 95.4 | 0.4 | <0.001 |
| CI | 1.30 | 0.14 | 1.31 | 0.20 | 0.52 |
| HI | 0.14 | 0.19 | 0.15 | 0.19 | 0.009 |
| <i>PTV2only (60 Gy)</i> | | | | | |
| Mean | 103.3 | 1.0 | 103.1 | 1.5 | 0.16 |
| D2 | 109.0 | 1.4 | 111.8 | 2.5 | 0.02 |
| D98 | 96.9 | 2.0 | 95.2 | 1.6 | 0.02 |
| CI* | 1.85 | 1.82 | 1.84 | 1.78 | 0.70 |
| HI | 0.12 | 0.03 | 0.16 | 0.03 | <0.001 |
| <i>PTV3only (50 Gy)</i> | | | | | |
| Mean | 103.7 | 2.4 | 103.0 | 2.9 | 0.005 |
| D2 | 116.8 | 5.0 | 120.3 | 5.7 | 0.005 |
| D98 | 98.6 | 5.1 | 96.4 | 5.6 | <0.001 |
| CI* | 1.71 | 0.72 | 1.74 | 0.72 | 0.01 |
| HI | 0.23 | 0.16 | 0.28 | 0.14 | <0.001 |

All doses are normalised to prescription dose (68, 66, 60 or 50 Gy).

* Conformity index calculated from the full PTV.

Table 4
Organ at risk mean doses.

| OAR | Unit | Autoplan | | Manual | | p |
|---------------------------|------|----------|------|--------|------|--------|
| | | Mean | STD | Mean | STD | |
| Spinal cord | (Gy) | 20.2 | 6.9 | 22.9 | 5.8 | <0.001 |
| Brainstem | (Gy) | 3.5 | 4.0 | 5.1 | 4.7 | <0.001 |
| Oral cavity | (Gy) | 31.6 | 13.3 | 34.3 | 12.8 | <0.001 |
| Lips of mouth | (Gy) | 12.3 | 7.7 | 15.2 | 6.8 | <0.001 |
| Parotid gland ipsi* | (Gy) | 23.4 | 16.4 | 25.5 | 15.7 | <0.001 |
| Parotid gland con** | (Gy) | 18.5 | 8.1 | 20.5 | 8.8 | 0.004 |
| Submandibular gland ipsi* | (Gy) | 53.2 | 11.4 | 56.0 | 7.7 | 0.01 |
| Submandibular gland con** | (Gy) | 34.0 | 19.2 | 40.5 | 18.9 | <0.001 |
| Mandible | (Gy) | 30.2 | 9.4 | 32.3 | 8.9 | <0.001 |
| Thyroid gland | (Gy) | 34.6 | 13.3 | 37.1 | 11.2 | <0.001 |
| Larynx | (Gy) | 39.1 | 9.4 | 44.8 | 8.7 | <0.001 |
| Body | (Gy) | 9.3 | 3.0 | 9.8 | 2.9 | <0.001 |

Mean doses of the patient specific mean OAR doses. AU reduced the OAR mean doses for all organs including the whole scanned part of the patient (Body).

* Ipsilateral gland.

** Contralateral gland.

is, the plan included a relatively large volume of lung tissue, which was irradiated to a lesser degree by the manually created plan.

Compared to the MA plans, a larger degree of MLC modulation was observed for the AU plans which are reflected in an increased number of monitor units in the AU plans. The pass rate was 97.7% in AU plans compared with 98.4% for MA plans assessed by ArcCheck measurements (Table 5). The average beam-on time of was 4 s longer in the AU plans. Mean operator time spent on MA plans was more than twice that of AU plans.

Three of the 30 patients had critical OAR adjacent to the PTV1 (high dose target) which required more operator time for both the MA and AU plans resulting in a relatively large variation in operator time.

4. Discussion

In 29 of the 30 patients the AU plan was selected. In most cases, the plan selection was straightforward with the AU plans clearly

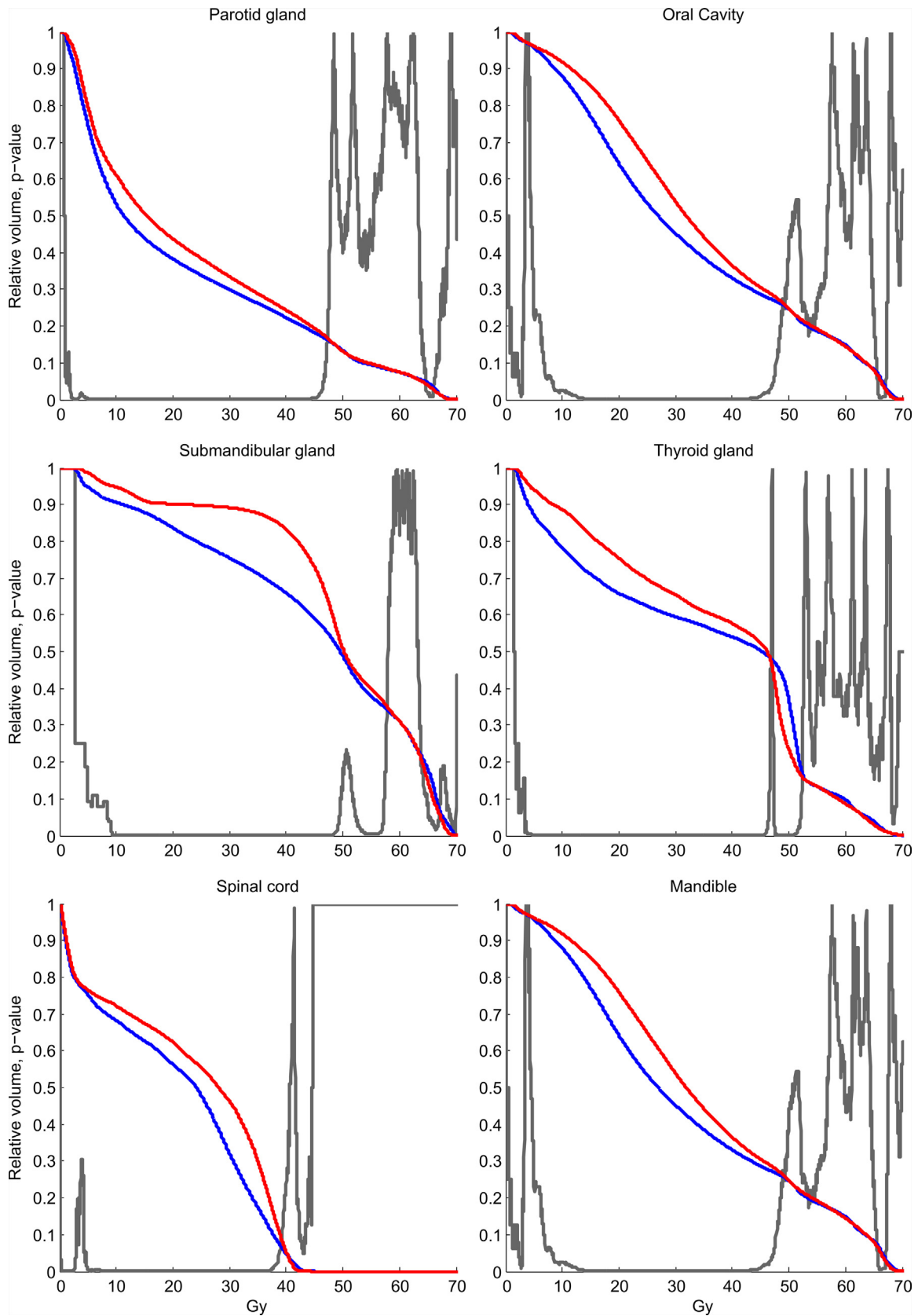


Fig. 1. Mean DVH of six OAR for all 30 patients. Red line is MA plan and blue line is AU plan. The grey p -value curve illustrates dose regions from around 10 to 45 Gy where AU plans significantly spare the organs more. The mean dose reductions for these six organs are between 2 to 5 Gy. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

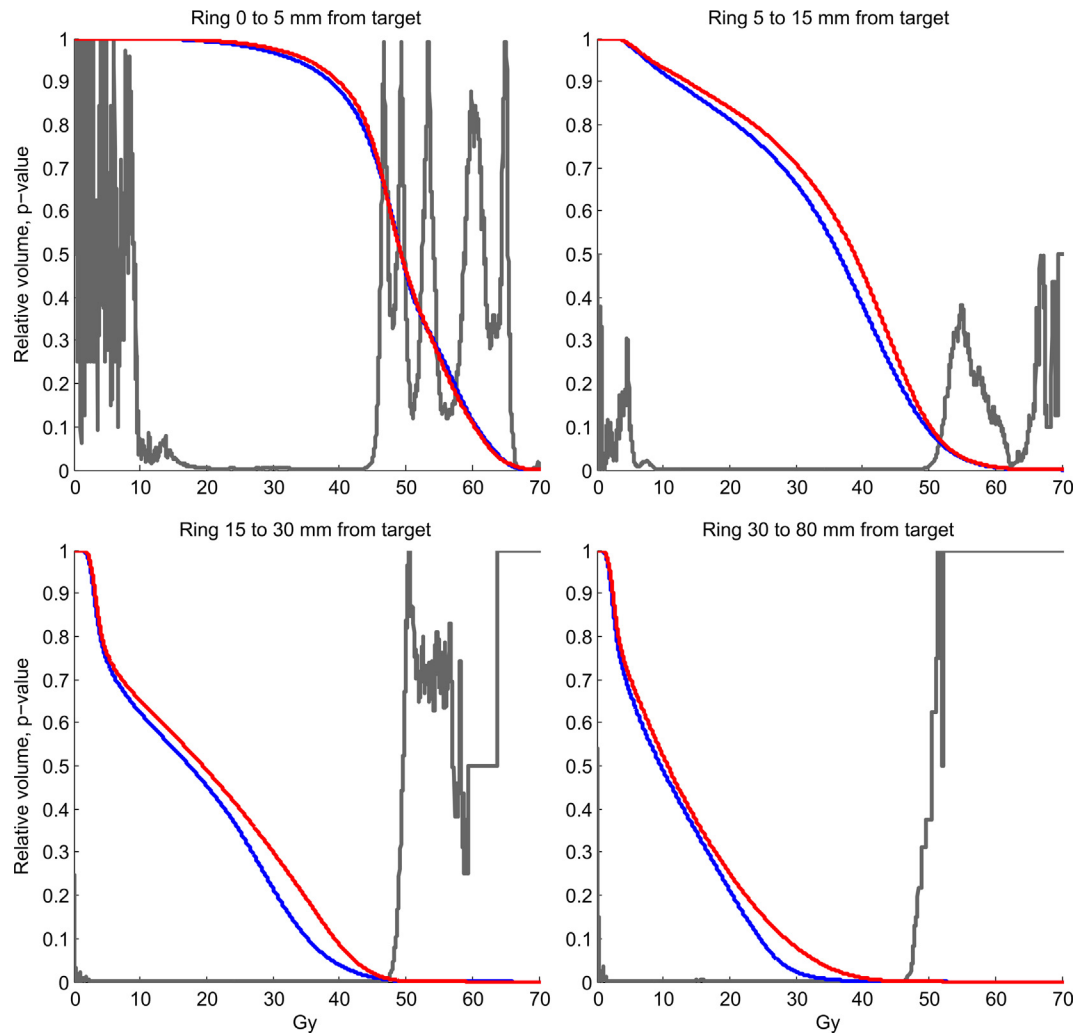


Fig. 2. Mean DVH of four ring-ROI's for all 30 patients. Black line is MA plan and dotted line is AU plan. Red line is MA plan and blue line is AU plan. The grey p -value curve illustrates where the AU and MA plans are significantly different.

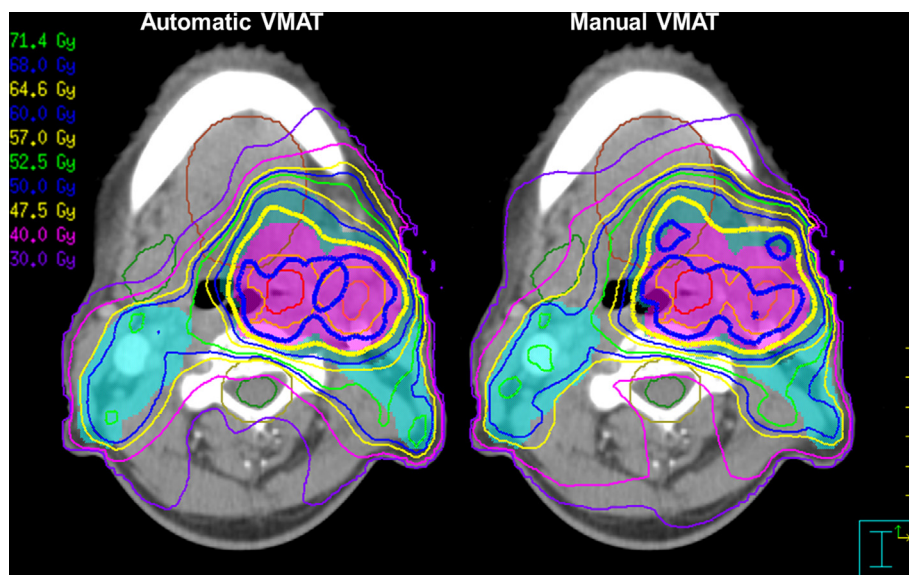


Fig. 3. Screen dump comparison of a representative patient. Iso-dose curves are 105% (green), 100% (blue) and 95% (yellow) for the three dose prescriptions (68 Gy, 60 Gy and 50 Gy). Dose spillage iso-dose curve are 40 and 30 Gy. The GTV tumour and GTV lymph node are shown in red and orange. CTV1 is shown in light orange, with corresponding PTV1 in colourwash purple. Colourwash blue and light blue represent PTV2 and PTV3. The contour of the extended oral cavity is brown. The right submandibular gland and spinal cord are shown in green and the PRV spinal cord in light brown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Treatment delivery metrics.

| Target | Unit | Autoplan | | Manual | | p |
|---------------------|------|----------|-----|--------|-----|--------|
| | | Mean | STD | Mean | STD | |
| MU | MU | 435 | 79 | 360 | 56 | <0.001 |
| Operator time | min | 32 | 26 | 64 | 31 | <0.001 |
| Beam on time | s | 114 | 22 | 110 | 24 | 0.017 |
| Pass rate (3%,3 mm) | % | 97.7 | 1.7 | 98.4 | 1.7 | 0.001 |
| Pass rate (2%,2 mm) | % | 90.6 | 4.3 | 92.5 | 4.5 | 0.003 |

Treatment delivery metrics when measured on an ArcCheck phantom.

sparing OAR better without compromising the target coverage. In a few cases the mean OAR doses were similar, with the AU plans only slightly lower. The only manual plan selected for clinical use was a non-standard H&N cancer involving level VI and thus including a large part of the lung. This indicates that the AU plan is more appropriate for well-established protocols and that special cases might require several runs with AU to obtain proper setting.

Previously, a research version of Autoplan has been tested in a retrospective study creating AU ssIMRT plans for 26 H&N patients [12]. That study investigated the potential of an Autoplan workflow and thus did not allow for a manual post optimisation of the AU plans. Similar to the current study, the AU plans improved the sparing of the OAR and had target coverage similar to the manual plans. However, due to the decision of no post optimisation of the AU, ssIMRT plans had slightly reduced target coverage relative to the manual plans. In the current study target dosages were manually fine-tuned during the manual post optimisation.

Krayenbuehl et al. showed in a retrospective study that Autoplan also can improve plan quality when compared to a different treatment planning system, Eclipse. This suggests that the plan quality improvement presented here is more generic and not solely related to Pinnacle [11]. Krayenbuehl et al. did not use any post AU optimisation, which explains the difference in planning time. They only spent on average 4 min setting up the AU plan, where in the current study the planning time included plan evaluation and post-AU optimisation.

In the post optimisation process extra objectives were added if violations of the guidelines for critical structures (spinal cord or brain stem) or targets were observed. Even with the critical structures set to “none compromise” the max dose could exceed the tolerance slightly for patients with targets close to the OAR. This was solved by adding a max dose objective with high weight for the specific OAR in the post optimisation. For the targets a single uniform dose objective was added for the CTV1 target (high dose target). This was done to reduce the mean CTV1 dose to match the prescription dose within $\pm 1\%$, as recommended within the DAHANCA guidelines. The “need” of an additional constraint on the high dose volume seems related to the difference between dose prescription as recommended by the ICRU and RTOG guidelines [15]. The ICRU guideline aims for a median target dose close to the prescribed dose and thereby allows the minimum dose (defined as D98%) to be 95% of the prescribed dose., RTOG aims for a minimum target dose (defined as D95%) equal to the prescribed dose. Therefore, the RTOG prescription will always give a higher mean and median dose than the ICRU prescription. Autoplan in Pinnacle seems to adhere mainly to the RTOG guidelines which could be the cause for an additional post Autoplan optimisation constraint on the high dose volume needed.

Automatic treatment planning is typically based on one of two approaches. One being the atlas based model, where a group of representative plans is used as a base for a model. A new treatment plan is then created by comparing targets and OAR anatomy with the model and from there derived the DVH objectives. One exam-

ple of this is the commercial RapidPlan used in the treatment planning system Eclipse [8,16]. The second approach is template based automatic treatment planning, where the automatic plan algorithm tries to drive the optimisation according to template originated from a protocol. Two such systems are the Erasmus-iCycle from the Rotterdam group which works with TPS Monaco [10] and the Autoplan from Pinnacle³ used in this study. Both approaches seem to create better or at least as good plans as a highly experienced dosimetrist. An additional benefit of the template based systems is that knowledge-sharing between centres is easier and more robust. A centre would be able to create plans of the same quality as presented in this study by coping Table 2. If their prescription doses or OAR dose limits are different they would just need to be adjusted in the AU template.

For the atlas based systems the plans quality will depend heavily on the quality of the plans in the atlas. With sparse resources at most centres it can be difficult to allocate time to generate an atlas of very high plan quality. Potentially the Autoplan generated plans could be used as the atlas database improving plan quality across different TPS.

One of the main benefits of AU plans is the general sparing of OAR compared to the MA plans. For H&N cancer treatment there has been a focus during the last years for deescalating the treatment doses for HPV positive tumours [17,18]. This is desired because the treatment dose could be lowered and hence the toxicity could be reduced and potentially improving the quality of life for the patients. With Autoplan the dose reduction to the OAR is around 10%, which is roughly the same dose reduction the HPV positive de-escalation trials are aiming for. The benefit of the Autoplan, however, is for all H&N cancer patients and comes with no higher risk of recurrences, since the target doses are maintained, hence, AU can be perceived as “free of charge”.

Another benefit of AU planning is in the adaptive radiotherapy setting. In this study replanning, during treatment, was necessary for 3 out of the 30 patients. Autoplan was also used to create the new plans; however they were not included in the analysis. The plans showed similar sparing of OAR and were substantially quicker to produce compared to conventional manual replans.

The AU plans achieve improved planned dose distributions, compared to MA, by adding more MLC modulation which can be observed by an average increase of 75 MU per plan in this study. This propagates into a slightly lower pass rate measure on the ArcCheck phantom, however all plans were clinically acceptable. Comparison to other TPS (Monaco or Eclipse) the number of MU used by Autoplan for head and neck is still low [19–21].

In Autoplan the prioritisation of OAR only has four levels (low, medium, high, non-compromise), which for some specific patients will be too low and therefore it can be difficult to find the appropriate compromise between target and multiple OARs. The automatic nature of Autoplan makes manual interaction time-consuming and therefore, a generic template is of vital importance.

5. Conclusion

In conclusion, the Autoplan-generated treatment plans were selected in 29 of 30 cases for clinical treatment of head and neck cancer when selected blindly head-on with the standard manual planning method. The target coverage was similar with maintained treatment doses while average OAR mean doses were reduced between 0.5 and 6.5 Gy. The manual time spent on planning was reduced by a factor of two and the delivery of the AU plans was of high clinical quality, though the pass rate was slightly lower compared to MA plans. From these results, AU planning has now been implemented as the clinical standard of head and neck treatment in our centre.

Authors' contributions

CRH and AB collected data, performed the data analysis and drafted the manuscript. IH helped create plans and collect the data. RZ, NG, JJ and JGE did the blinded clinical evaluation of the manual and automated optimised plans. CRH, AB and CB conceived the study. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work is related to AgeCare (Academy of Geriatric Cancer Research), an international research collaboration based at Odense University Hospital, Denmark.

References

- [1] 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(18):2996–3001.
- [2] Clark CH, Aird EG, Bolton S, Miles EA, Nisbet A, Snaith JA, et al. Radiotherapy dosimetry audit: three decades of improving standards and accuracy in UK clinical practice and trials. *Br J Radiol* 2015;88(1055):20150251.
- [3] Christianen MEMC, Schilstra C, Beetz I, Muijs CT, Chouvalova O, Burlage FR, et al. Predictive modelling for swallowing dysfunction after primary (chemo) radiation: results of a prospective observational study. *Radiother Oncol* 2012;105(1):107–14.
- [4] Brouwer CL, Steenbakkers RJHM, Gort E, Kamphuis ME, van der Laan HP, Van'tVeld AA, et al. Differences in delineation guidelines for head and neck cancer result in inconsistent reported dose and corresponding NTCP. *Radiother Oncol* 2014;111(1):148–52.
- [5] Bertelsen A, Hansen CR, Johansen J, Brink C. Single arc volumetric modulated arc therapy of head and neck cancer. *Radiother Oncol* 2010;95(2):142–8.
- [6] Fiorino C, Dell'Oca I, Pierelli A, Broggi S, De Martin E, Di Muzio N, et al. Significant improvement in normal tissue sparing and target coverage for head and neck cancer by means of helical tomotherapy. *Radiother Oncol* 2006;78(3):276–82.
- [7] Simone 2nd CB, Ly D, Dan TD, Ondos J, Ning H, Belard A, et al. Comparison of intensity-modulated radiotherapy, adaptive radiotherapy, proton radiotherapy, and adaptive proton radiotherapy for treatment of locally advanced head and neck cancer. *Radiother Oncol* 2011;101(3):376–82.
- [8] 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(3):385–91.
- [9] Ruan D, Shao W, Demarco J, Tenn S, King C, Low D, et al. Evolving treatment plan quality criteria from institution-specific experience. *Med Phys* 2012;39(5):2708–12.
- [10] Voet PWJ, Breedveld S, Dirkx MLP, Levendag PC, Heijmen BJM. Integrated multicriterial optimization of beam angles and intensity profiles for coplanar and noncoplanar head and neck IMRT and implications for VMAT. *Med Phys* 2015;42(1):226–35.
- [11] Krayenbuehl J, Norton I, Studer G, Guckenberger M. Evaluation of an automated knowledge based treatment planning system for head and neck. *Radiat Oncol* 2015;10(1):226.
- [12] Hazell I, Bzdusek K, Kumar P, Hansen CR, Bertelsen A, Eriksen JG, et al. Automatic planning of head and neck treatment plans. *J Appl Clin Med Phys* 2016;17(1).
- [13] DAHANCA radiotherapy guidelines 2013 English 2.0 Jan 2015.pdf, <https://www.dahanca.oncology.dk/Brows_Web_Guidelines>; [accessed 01.05.2016].
- [14] ICRU 83 prescribing, recording, and reporting photon-beam IMRT Technical report. 7010 Woodmont Avenue Bethesda, Maryland 20814, U.S. A.: International Commission on Radiation Units and Measurements; 2010.
- [15] RTOG RTQA protocol prescription guidelines, <<https://www.rtog.org/CoreLab/RTQAProtocolPrescriptionGuidelines.aspx>>; [accessed 01.05.2016].
- [16] Tol JP, Delaney AR, Dahele M, Slotman BJ, Verbakel WFAR. Evaluation of a knowledge-based planning solution for head and neck cancer. *Int J Radiat Oncol Biol Phys* 2015;91(3):612–20.
- [17] Vokes EE, Agrawal N, Seiwert TY. Hpv-associated head and neck cancer. *J Natl Cancer Inst* 2015;107(12).
- [18] Mirghani H, Amen F, Blanchard P, Moreau F, Guigay J, Hartl DM, et al. Treatment de-escalation in hpv-positive oropharyngeal carcinoma: ongoing trials, critical issues and perspectives. *Int J Cancer* 2015;136(7):1494–503.
- [19] Vanetti E, Clivio A, Nicolini G, Fogliata A, Ghosh-Laskar S, Agarwal JP, et al. Volumetric modulated arc radiotherapy for carcinomas of the oro-pharynx, hypo-pharynx and larynx: a treatment planning comparison with fixed field IMRT. *Radiother Oncol* 2009;92(1):111–7.
- [20] Wiezorek T, Brachwitz T, Georg D, Blank E, Fotina I, Habl G, et al. Rotational IMRT techniques compared to fixed gantry IMRT and tomotherapy: multi-institutional planning study for head-and-neck cases. *Radiat Oncol* 2011;6:20.
- [21] Kantz S, Söhn M, Troeller A, Reiner M, Weingandt H, Alber M, et al. Impact of MLC properties and IMRT technique in meningioma and head-and-neck treatments. *Radiat Oncol* 2015;10:184.