

Automation of radiation treatment planning

Evaluation of head and neck cancer patient plans created by the Pinnacle³ scripting and Auto-Planning functions

Stefan Speer¹ · Andreas Klein² · Lukas Kober³ · Alexander Weiss¹ · Indra Yohannes⁴ · Christoph Bert¹

Received: 8 July 2016 / Accepted: 10 May 2017 / Published online: 26 June 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract

Background Intensity-modulated radiotherapy (IMRT) techniques are now standard practice. IMRT or volumetric-modulated arc therapy (VMAT) allow treatment of the tumor while simultaneously sparing organs at risk. Nevertheless, treatment plan quality still depends on the physicist's individual skills, experiences, and personal preferences. It would therefore be advantageous to automate the planning process. This possibility is offered by the Pinnacle³ treatment planning system (Philips Healthcare, Hamburg, Germany) via its scripting language or Auto-Planning (AP) module.

Materials and methods AP module results were compared to in-house scripts and manually optimized treatment plans for standard head and neck cancer plans. Multiple treatment parameters were scored to judge plan quality (100 points = optimum plan). Patients were initially planned manually by different physicists and re-planned using scripts or AP.

Results and discussion Script-based head and neck plans achieved a mean of 67.0 points and were, on average,

superior to manually created (59.1 points) and AP plans (62.3 points). Moreover, they are characterized by reproducibility and lower standard deviation of treatment parameters. Even less experienced staff are able to create at least a good starting point for further optimization in a short time. However, for particular plans, experienced planners perform even better than scripts or AP. Experienced-user input is needed when setting up scripts or AP templates for the first time. Moreover, some minor drawbacks exist, such as the increase of monitor units (+35.5% for scripted plans).

Conclusion On average, automatically created plans are superior to manually created treatment plans. For particular plans, experienced physicists were able to perform better than scripts or AP; thus, the benefit is greatest when time is short or staff inexperienced.

Keywords Radiotherapy, intensity-modulated · Organs at risk · Spinal cord · Brainstem · Parotid gland

Automatisierte Bestrahlungsplanung

Auswertung von mit Pinnacle³ via Scripting und Auto-Planning erzeugten Bestrahlungsplänen von Patienten mit Kopf-Hals-Tumor

Zusammenfassung

Hintergrund Intensitätsmodulierte Strahlentherapie (IMRT) hat sich als Standard durchgesetzt. Mit IMRT oder volumenmodulierter Arc-Therapie (VMAT) lassen sich Tumore behandeln und gleichzeitig die Risikoorgane schonen. Dennoch ist die Planungsqualität abhängig von individuellem Geschick, Erfahrung und persönlichen Vorlieben des Planers. Vorteilhaft wäre daher ein automatisierter Planungsprozess. Das Bestrahlungsplanungssystem Pinnacle³

✉ Dr. Stefan Speer
stefan.speer@uk-erlangen.de

¹ Department of Radiation Oncology, Universitätsklinikum Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Universitätsstraße 27, 91054 Erlangen, Germany

² EKS Engineering GmbH, Dr.-Mack-Straße 88, 90762 Fürth, Germany

³ Strahlentherapie Tauber-Franken, Uhlandstraße 7, 97980 Bad Mergentheim, Germany

⁴ Rinecker Proton Therapy Center, Schäftlarnstraße 133, 81371 Munich, Germany

(Philips Healthcare, Hamburg, Deutschland) ermöglicht dies mit einer internen Scripting-Funktion oder dem Auto-Planning-(AP-)Modul.

Material und Methoden Die Ergebnisse des AP-Moduls wurden mit denen von intern erstellten Skripten und manuellen Bestrahlungsplänen verglichen. Um die Planqualität zu beurteilen, wurden mehrere Behandlungsparameter bewertet (optimaler Plan = 100 Punkte). Alle Patienten wurden initial von verschiedenen Physikern manuell und anschließend scriptbasiert oder mit AP erneut geplant.

Ergebnisse Skriptbasierte HNO-Pläne erzielten im Mittel 67,0 Punkte und waren manuell geplanten (59,1 Punkte) und AP-Fällen (62,3 Punkte) im Durchschnitt überlegen. Zudem zeigten sie eine hohe Reproduzierbarkeit und damit einhergehend eine geringere Standardabweichung der Planparameter. Sogar weniger erfahrenem Personal ist es möglich, bei geringerem Zeitbedarf zumindest einen guten Ausgangspunkt für weitere Nachoptimierungen zu erzeugen. Erfahrene Planer schneiden allerdings bei einigen Plänen besser ab als Skript oder AP. Beim ersten Erstellen von Skripten oder AP-Templates sind Vorgaben von erfahrenen Nutzern nötig. Weiterer Nachteil ist z. B. ein Anstieg der Monitoreinheiten (Skript-Pläne +35,5%).

Schlussfolgerung Insgesamt ist die automatische Planung der manuellen im Durchschnitt überlegen. Einzelne Pläne gelingen erfahrenem Personal besser als Skript oder AP. Der größte Nutzen besteht somit bei Zeitmangel und unerfahrenem Personal.

Schlüsselwörter Intensitätsmodulierte Strahlentherapie · Risikoorgane · Rückenmark · Hirnstamm · Parotis

Nowadays, intensity-modulated radiotherapy (IMRT) or volumetric-modulated arc therapy (VMAT) techniques are standard practice in radiation therapy. Nevertheless, treatment plan quality still depends on the experience of the planner, or may even differ due to personal preferences. In addition, the time needed for plan creation may prevent an adequate plan output.

Several options exist to overcome these limitations, such as multicriteria optimization (MCO) [1, 2], knowledge-based treatment planning [3], or automated treatment planning [4–6].

Other solutions can be constructed using the scripting tool provided by the Pinnacle³ treatment planning system (Philips Healthcare GmbH, Hamburg, Germany) or via python in the RayStation (RaySearch Laboratories AB, Stockholm, Sweden). In this work, Pinnacle³ was used.

The basic scripting feature records frequent steps of the planning process. Once recorded, these scripts/macros are able to speed up the process tremendously. Such Pinnacle³ scripts can be manually extended for more complex tasks. A similar way to generate treatment plans is Pinnacle³'s

Auto-Planning (AP) module, which can be considered to be a commercial “scripting solution”.

Within the scope of this work, scripts were created that trace in-house methodology of generating (VMAT) head and neck cancer treatment plans. Accordingly, an AP template was created. The goal of this work is to provide information about the quality of treatment plans derived by scripting and those generated with the AP module, each in comparison to manually created treatment plans. For evaluation, a score based on a set of parameters (according to International Commission on Radiation Units and Measurements, ICRU, report 83) or others especially determined for this purpose was generated (see Table 3). The time needed for treatment planning was also analyzed.

Materials and methods

Implementation of scripted planning and Auto-Planning

In most facilities, the preparation of a treatment plan follows initial steps such as definition of beam parameters and other customization. These frequent steps were recorded with Pinnacle³'s scripting tool. The resulting script is written in Pinnacle³'s scripting language, whose basic instructions can be adapted to one's own needs. For more information, the very instructive summary by Sean Geoghegan [7] can be recommended.

Treatment plan optimization starts with a scripted check of whether the mandatory organs at risk (OAR) are present and drawn correctly to ensure proper functioning. These checks include anatomic consistency (e. g., seamless transition from spinal cord to brainstem).

The optimization process was configured by arranging dose objectives via scripts in Pinnacle³'s classic IMRT dose optimization window or in the AP dose goal window, depending on planning technique. To achieve a clinically acceptable dose distribution, additional artificial structures are needed. This applies especially to the classic procedure of IMRT dose optimization, which requires a lot of experience and recurrent operation. The AP function is more user friendly and requires less input to reach the predefined goals. However, the experience of the planner is still a considerable factor. Table 1 lists the main and auxiliary structures with the implemented objectives for classic IMRT planning (scripted) and AP.

Script and AP goals are configured to generate a planning target volume (PTV) dose of 50 Gy (25 fractions of 2 Gy). Potential subsequent boost series were neither scripted nor evaluated. This approach is adapted to the workflow at the Department of Radiation Oncology, University Hospital Erlangen. Since the tumor may change considerably in size, the boost volume often cannot be defined adequately at the

start of treatment and the conception is thus to decline simultaneous integrated boost techniques (SIB) in most cases. Nevertheless, scripts and AP could easily be adapted to SIB procedures.

The goals listed in Table 1 are more or less equal to the classic IMRT objectives, with the exception that AP does not need most of the auxiliary structures. AP automatically detects structures overlapping with the target and creates internal structures and settings according to the “compromise” option. As recommended at the Pinnacle³ AP user meeting, two goals were given to the optimizer for a certain OAR whenever a strong dose gradient is desired. In this case the “max dose” and “mean dose” arguments were passed to Pinnacle³.

AP has problems limiting the maximum dose of a structure lying at least partially inside the PTV, because disabling the “compromise” option can be considered as being too strong for the desired behavior, whereas using the “compromise” option obviously ignores the maximum dose inside the PTV. However, the resulting hotspots can be removed by a manual or scripted re-optimization after AP has finished. In the presented data, no such re-optimizations took place.

For scripted and manual IMRT plans, a single optimization cycle will usually not lead to a satisfying dose distribution and a new iteration cycle has to be started. A typical automated procedure has been published by Xhaferllari et al. [8], whereby their approach focusses on in-field boost techniques and does not deal with AP at all. Therefore, the

Table 1 Initial setup: required structures, their dose objectives, and weight/priority as used for optimizing treatment plans for Pinnacle³ (Philips Healthcare, Hamburg, Germany) scripting/Auto-Planning (AP). For AP, the “compromise option” is activated for every contour except spinal cord and brainstem

Contour name	Explanation	Classic IMRT		AP	
		Dose objective	Weight	Dose objective	Priority
PTV plan	PTV without skin (4 mm), but including GTV or lymph nodes if applicable	50 Gy uniform	20	50 Gy (target)	–
PTV part par left/right (“local parotid coverage”)	Part of the PTV close to the parotid to prevent dose intrusions (parotid plus 10 mm, avoiding PTV minus 10 mm lateral and exterior PTV)	50 Gy/99% min	2.5	–	–
GTV/lymph nodes	Only if structures are present	50.5 Gy/100% min	1.5	–	–
Spinal cord	AP: compromise option <i>not</i> activated	–	–	31.94 Gy max	High
Spinal cord +4 mm	Plus 4 mm to account for uncertainties	27 Gy max	5	–	–
Brainstem	AP: compromise option <i>not</i> activated	–	–	37 Gy max	High
Brainstem +4 mm	Plus 4 mm to account for uncertainties	34 Gy max	1	–	–
Eye left/right	–	13.5 Gy max	0.01	15 Gy max	Med
Lens left/right	Looking cranial during irradiation could lower dose in most cases	1.4 Gy max	0.01	2 Gy max	Med
Optic system	Union of optic nerves and chiasm plus 5 mm to account for uncertainties	48.5 Gy max	0.01	30 Gy max	Med
Optic system without PTV	Above contour excluding PTV plus 6 mm	14.25 Gy max	0.01	–	–
Brain	Complete organ	–	–	30 Gy max	Low
Brain without PTV	Only parts excluding PTV plus 6 mm	28.5 Gy max	0.05	–	–
Mandible	–	48 Gy max	2	–	–
Oral cavity	If GTV or lymph nodes are located near or inside oral cavity, the contour is shrunk respectively	48 Gy max	2	–	–
Acoustic N./cochlea left/right	Union of acoustic nerve and cochlea plus 5 mm to account for uncertainties	49.5 Gy max	0.2	30 Gy max	Low
Acoustic N./cochlea left/right without PTV	Above contour excluding PTV plus 6 mm	24.5 Gy/2% max	1	–	–
Ring	Circular structure between PTV plus 6 mm and 26 mm, plus dorsal extension to prevent hotspots in the neck region	29 Gy/10% max	0.1	–	–
Parotid left/right	–	48.5 Gy max	1	10 Gy max + 10 Gy mean	Med Med

Table 1 Initial setup: required structures, their dose objectives, and weight/priority as used for optimizing treatment plans for Pinnacle³ (Philips Healthcare, Hamburg, Germany) scripting/Auto-Planning (AP). For AP, the “compromise option” is activated for every contour except spinal cord and brainstem (Continued)

Contour name	Explanation	Classic IMRT		AP	
		Dose objective	Weight	Dose objective	Priority
Parotid left/right without PTV	Parotid excluding PTV plus 5 mm	14 Gy/4% max	1.5	–	–
Lung left/right	Complete organ	–	–	20 Gy max	Low
Lung without PTV	Both lungs excluding PTV plus 26 mm	19 Gy/10% max	0.2	–	–
Artifacts	Area of metal artifacts (if present) excluding PTV plus 10 mm	20 Gy/10% max	0.01	25 Gy max	Low
Dors. esophagus	Auxiliary structure typically located near the esophagus excluding PTV plus 5 mm	28.5 Gy max	0.3	30 Gy	Med
Larynx 1	Area of spared larynx excluding PTV plus 5 mm	26.5 Gy max	2	–	–
Larynx 2	Larynx 1 plus 10 mm	–	–	10 Gy max + 10 Gy mean	High High
Lips	Complete organ (only if completely covered by PTV plus 6 mm)	48 Gy max	0.01	25 Gy max	Med
Lips without PTV	Lips excluding PTV plus 6 mm (only if lips are not completely covered by PTV plus 6 mm)	20 Gy max	0.05	–	–
Mouth (zone 1,2,3)	Mouth, divided into consecutive ventral areas (6, 10, and 26 mm from PTV)	47.5 Gy max	0.03	–	–
		40 Gy max	0.03	–	–
		30 Gy max	2	30 Gy max	Med
Normal tissue	Tissue excluding PTV plus 26 mm	38.5 Gy max	0.02	–	–
Skin	Skin (thickness 3 mm) excluding PTV	40 Gy max	1	–	–
Skin	Above contour plus 5 mm without air and PTV	–	–	47.5 Gy max	Med
Missing Contour	“Dummy” will be inserted instead (PTV/uniform dose)	50 Gy uniform	0.5	–	–

PTV planning target volume, PTV part par PTV partial/parotid gland (“local coverage”), GTV gross tumor volume, min minimum, max maximum, Med medium, N. nerve, Dors. dorsal

Table 2 Auxiliary structures in between optimization cycles

Contour	Explanation	Dose Objective	Weight
Maximum	Contour of 107% isodose line	52.5 Gy max	0.03
Cold spot	Parts of PTV where dose falls below 95%	49.5 Gy/99% min	1.5
Margin	Contour of 95% isodose line outside PTV	47.5 Gy/5% max	0.2
80% mod	Contour of 80% isodose line excluding PTV plus 12 mm	39 Gy	0.05
60% mod	Contour of 60% isodose line where located inside or close to brain, brainstem, spinal cord, and auxiliary contours “dors. esophagus,” “larynx,” and “zone 3”	29 Gy	0.1

PTV planning target volume, min minimum, max maximum, mod modified, dors. dorsal

dose objectives were adapted to the local experience gathered from manually created VMAT head and neck treatment plans.

Table 2 lists the structures and dose constraints that are automatically created by the scripts before a new optimization cycle is initialized. After cycle three, the structures of Table 2 are complemented by further typical auxiliary contours, partially depending on the individual optimization

progress. The maximum number of cycles is limited to six, because no relevant changes are expected afterwards.

Evaluation of strategy

In order to decide whether the plan is clinically acceptable, a set of parameters exceeding typical values such as coverage were defined (see Table 3).

Table 3 Parameters for treatment quality benchmarking. Minimum and maximum scores are given. Values between are interpolated linearly

Parameter	Explanation	Score (points)
Coverage C	$C = \frac{V(\text{PTV covered by 95\%})}{V_{\text{PTV}}}$	90% (0 points) < C < 95% (20 points)
Conformity	$K = \frac{V(\text{PTV covered by 95\%})^2}{V(95\%) \cdot V_{\text{PTV}}}$	0.5 (0 points) < K < 1 (10 points)
Min./max. dose	D (Vol.PTV = 98%), D (Vol.PTV = 2%), D_{max} (2 cm ³) [9]	51 Gy (5 points) < D_{max} < 55 Gy (0 points)
Mean dose	D_{mean} and D_{50} respectively	–
Homogeneity H	$H = \frac{D(\text{Vol.PTV}=2\%) - D(\text{Vol.PTV}=98\%)}{D_{50}}$	0.1 (10 points) < H < 0.2 (0 points)
Dose gradient G_1, G_2	$G_x = \frac{D_{\text{mean, Ring } x}}{D_{\text{mean, PTV}}}$ Ring 1 is a circular structure from 3 mm to 10 mm from PTV; ring 2 from 10 mm to 30 mm from PTV (without air) [10]	–
Parotid gland	D_{mean} of parotid gland	10 Gy (5 points) < D_{mean} < 40 Gy (0 points)
Coverage in parotid gland area	C_{part} coverage (95% of reference dose) of auxiliary contour “PTV part par left/right”	0% (0 points) < C_{part} < 100% (5 points)
Larynx	Volume receiving > 30 Gy (auxiliary contour larynx)	–
Mandible	D_{max} (1 cm ³)	50 Gy (5 points) < D_{max} < 55 Gy (0 points)
Brain	Volume receiving > 30 Gy (outside PTV + 6 mm)	–
Oral cavity ($D_{0.5\text{cm}^3}$)	D_{max} (0.5 cm ³)	50 Gy (5 points) < D_{max} < 55 Gy (0 points)
Zone 3	Relative volume V receiving > 30 Gy of auxiliary contour “Zone 3”	0% (5 points) < V < 15% (0 points)
Acoustic nerves/cochlea	D_{max} (0.1 cm ³) and D_{mean} respectively of acoustic nerves/cochlea plus 6 mm	–
Spinal cord	D_{max} (0.5 cm ³)	27 Gy (5 points) < D_{max} < 31.5 Gy (0 points)
Brainstem	D_{max} (0.5 cm ³)	27 Gy (5 points) < D_{max} < 31.5 Gy (0 points)
Skin	Volume at 45/47.5 Gy, PTV excluded	2 cm ³ (5 points) < $V_{45\text{Gy}}$ < 15 cm ³ (0 points)
Optic system	D_{max} (only important if close to PTV)	–
Monitor units	Total amount of monitor units (MU) per 2 Gy	600 (5 points) < MU < 1000 (0 points)

PTV planning target volume, D_{max} maximum (point) dose, D_{mean} mean dose, D_{50} dose at 50% volume, PTV part par PTV partial/parotid gland (“local coverage”)

For example, a particular parotid mean dose is determined by its fraction covered by the PTV. However, the mean dose may be misinterpreted when it is achieved at the expense of missing (local) coverage. Most of these values can be quickly derived by using Pinnacle³'s score card. These parameters were scored according to Table 3, whereupon an optimum plan is considered to have 100 points.

Ten head and neck cancer patients were investigated. All patients were treated with a manually created VMAT plan during the period from October 2013 to August 2014. Patients who were not treated with VMAT or whose PTV dimensions would require a different collimator angle (due to the limitation of the accelerator's maximum aperture) were also rejected.

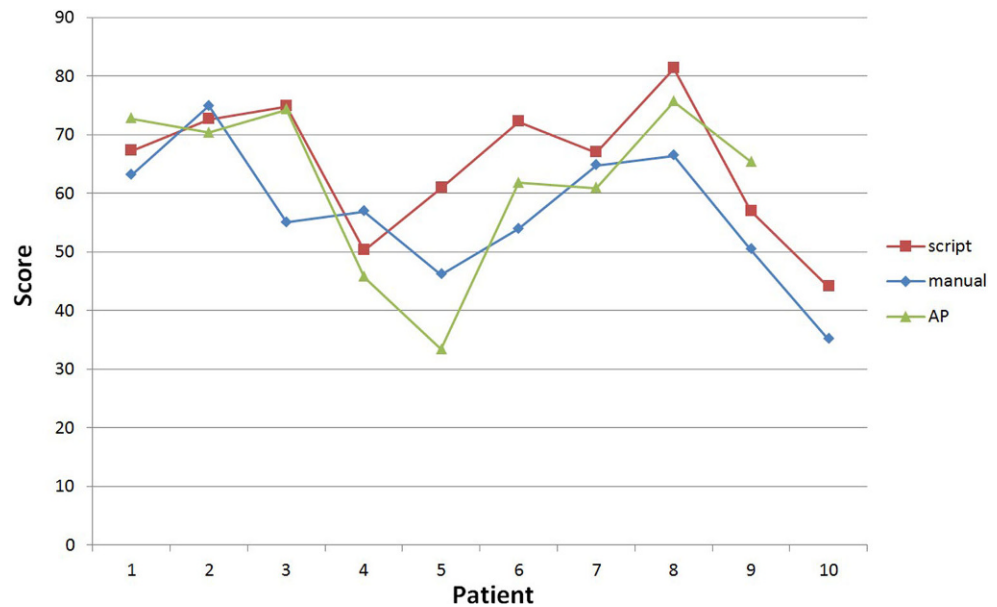
The manually created Pinnacle³ plans (version 9.4) comprise two arcs (200 to 160° and back, collimator 0°). These plans were generated by different planners (excluding plans

from the main author to avoid potential analogy to scripted plans) who had experience in treatment planning of at least 2 years. In the following, this set of plans will be labeled “manual.”

The automatically made plans are labeled “scripted” (internal script version 4.1) and “AP”. They were generated using Pinnacle³ version 9.10 (since version 9.4 does not provide the AP function) according to the methodology introduced above. For comparability, the same arc settings (200–160°, collimator 0°) were used in the analysis, whereas 181–180° arcs with collimator 10° would currently be used clinically. This approach results in a reduction of degrees of freedom, which influences plan quality and even caused AP to completely fail for one patient (rejected from evaluation).

The results were evaluated without any manual adjustment, except when the D_{50} (dose at 50% volume of PTV)

Fig. 1 Scores achieved by the manually created, scripted, and Auto-Planning module (AP; Philips Healthcare, Hamburg, Germany) treatment plans for the individual patients



differed more than 1% from 50 Gy. In this case, monitor units (MU) were adjusted accordingly. The time needed for treatment plan optimization was evaluated from a meaningful sample.

Results

When looking at the average score per treatment plan, scripted plans (67.0 points) were superior to AP (mean of 62.3 points) and manual plans (mean of 59.1 points). Nevertheless, when focusing on particular plans, experienced physicists were able to perform better than scripts or AP (Fig. 1). Obviously, the benefit is greatest when staff are short of time or unexperienced.

PTV-related parameters

Table 4 lists the detailed parameters as defined in Table 3 to benchmark the PTV coverage. As expected, coverage is better for scripted (94.1%) and AP (93.5 vs. 92.8%) plans, and associated with a lower standard deviation (1.6/2.2% vs. 2.4%).

Conformity (0.88/0.88/0.87) can be considered to be equal for all plans, whereas the homogeneity (0.12) of scripted plans seems to be slightly better than for other plans (0.14/0.14).

Normal tissue “hotspot” dose and some gradient parameters, G_1/G_2 , appear to be slightly better for manual plans but are of minor interest compared to the aforementioned parameters.

High-priority organs at risk

As mentioned above, the reference dose for the PTV is 50 Gy. Subsequent boost series (which are not subject to evaluation here) will raise the dose in high-risk regions to up to 72 Gy at most. Therefore, some dose toward the listed limits of brainstem and spinal cord has to be spared for potential dose boosting.

Thus, assuming a worst-case scenario and a final dose of 72 Gy, the dose delivered to the brainstem and the spinal cord should not exceed 37.5 or 31.9 Gy, respectively, at the end of the PTV irradiation. This refers to in-house accepted maximum (point) doses of 54 Gy for the brainstem and 46 Gy for the spinal cord (lower than Quantitative Analyses of Normal Tissue Effects in the Clinic, QUANTEC, doses [11]).

All plans—manual, scripted, and AP—usually fall below these point doses. Most variance can be observed when the brainstem is located very close to the PTV, because scripted plans in particular tend to have rather better coverage than OAR sparing. Table 4 provides detailed information; $D_{0.5\text{cm}^3}$ (dose at 0.5 cm³ volume) values are also given.

Judging parotid gland doses is more difficult, because the fraction that is covered by the PTV varies noticeably. In the cohort of patients studied, on average, 23% of the parotid gland volume is covered by the PTV and therefore the clinical dose limit (mean dose, D_{mean} , <25 Gy) [11] is very hard to achieve.

To compare differently located parotid glands, one could use the overlap volume histogram (OVH) method [12]; however, during analysis of the data, it was found that there is a nearly linear correlation between mean dose and the volume covered by the PTV (Fig. 2). This linear behavior

Table 4 Planning target volume (PTV) and organs at risk (OAR) parameters/score according to the definitions provided in Table 3

	Manual treatment plans				Scripted treatment plans				AP treatment plans			
	Min	Max	Mean	Dev	Min	Max	Mean	Dev	Min	Max	Mean	Dev
Coverage <i>C</i> [%]	87.50	95.59	92.83	2.35	91.75	96.12	94.14	1.59	89.06	95.53	93.50	2.22
Conformity <i>K</i>	0.85	0.91	0.88	0.02	0.86	0.90	0.88	0.02	0.86	0.89	0.87	0.01
<i>D</i> (<i>V</i> _{PTV=98%}), [Gy]	42.67	46.45	45.05	1.12	45.42	46.86	46.23	0.48	43.87	46.24	45.32	0.80
<i>D</i> (<i>V</i> _{PTV=2%}), [Gy]	51.45	53.47	52.08	0.58	52.14	52.89	52.42	0.23	51.88	53.12	52.41	0.41
Max. 2 cm ³ [Gy]	52.17	54.31	52.98	0.62	53.32	54.18	53.68	0.33	52.41	54.75	53.55	0.69
<i>D</i> _{mean} [Gy]	49.29	50.24	49.80	0.30	49.57	49.91	49.77	0.13	49.84	50.19	50.05	0.11
<i>D</i> ₅₀ [Gy]	49.83	50.39	50.11	0.22	49.73	50.01	49.87	0.11	50.20	50.48	50.35	0.09
Gradient <i>G</i> ₁	0.62	0.75	0.67	0.05	0.68	0.71	0.70	0.01	0.66	0.70	0.68	0.01
Gradient <i>G</i> ₂	0.33	0.46	0.39	0.04	0.38	0.43	0.40	0.01	0.36	0.41	0.37	0.01
Homogeneity <i>H</i>	0.11	0.18	0.14	0.02	0.11	0.14	0.12	0.01	0.12	0.18	0.14	0.02
Normal tissue [Gy]	26.20	39.03	34.25	3.93	36.94	39.76	38.66	0.96	33.48	47.78	39.83	4.46
Spinal Cord <i>D</i> _{max} [Gy]	29.01	32.27	31.13	0.95	27.19	30.51	28.18	0.95	30.62	35.08	32.67	1.45
Spinal Cord <i>D</i> _{0.5cm³} [Gy]	27.46	29.94	29.29	0.88	25.81	28.64	26.90	0.80	28.70	33.71	30.73	1.47
Brainstem <i>D</i> _{max} [Gy]	28.74	35.56	31.92	2.18	28.55	38.04	32.94	2.95	23.43	36.81	31.67	5.07
Brainstem <i>D</i> _{0.5cm³} [Gy]	26.54	31.44	28.88	1.64	25.31	30.41	28.89	1.66	18.17	32.62	25.36	5.48
Larynx <i>V</i> >30 Gy [cm ³]	11.47	52.42	24.72	17.32	4.36	21.05	10.28	7.32	12.71	34.03	25.62	8.88
Mandible <i>D</i> _{1cm³} [Gy]	48.97	52.14	50.69	0.99	50.10	52.78	51.34	0.82	50.35	52.74	51.46	0.67
Oral Cav. <i>D</i> _{0.5cm³} [Gy]	51.02	54.39	52.20	1.08	50.07	52.59	51.23	0.72	51.91	53.45	52.46	0.47
Brain >30 Gy [cm ³]	0.24	14.07	7.11	5.67	0.19	2.19	1.11	0.79	0.33	13.08	4.00	4.86
Mouth Zone 3 [%]	0.44	82.00	28.00	31.47	0.82	22.85	7.12	7.09	0.01	1.71	0.46	0.51
Acoustic <i>D</i> _{0.1cm³} [Gy]	40.09	51.32	48.55	3.54	35.17	51.34	47.70	5.19	28.46	51.12	44.55	7.09
Acoustic <i>D</i> _{mean} [Gy]	7.74	28.12	16.54	5.96	7.86	18.90	13.14	3.24	9.90	19.14	12.80	3.45
Skin <i>V</i> _{45Gy} [cm ³]	1.37	20.87	9.35	6.41	2.61	12.36	7.58	2.87	1.80	10.88	7.02	3.15
Skin <i>V</i> _{47.5Gy} [cm ³]	0.06	5.91	1.76	1.89	0.06	0.40	0.20	0.11	0.05	1.40	0.36	0.49
Monitor units	567	846	716.3	86.4	859	1080	970.9	70.4	486	837	723.8	104.9
Total Score/patient	46.16	74.91	59.07	8.96	50.30	81.28	67.03	9.62	33.39	75.62	62.26	14.23

AP plans created using the Pinnacle³ Auto-Planning module (Philips Healthcare, Hamburg, Germany), *Min* minimum, *Max* maximum, *Dev* standard deviation, *D*_{max} maximum dose, *D*_{mean} mean dose, *D*₅₀ dose at 50% volume, *D*_{0.1 cm³} dose at 0.1 cm³ volume, *D*_{0.5 cm³} dose at 0.5 cm³ volume, *D*_{1 cm³} dose at 1 cm³ volume, *V*_{45 Gy} volume at 45 Gy dose, *V*_{47.5 Gy} volume at 47.5 Gy dose

Fig. 2 Nearly linear correlation between mean dose of the parotid gland and its volume covered by the planning target volume (PTV). AP Auto-Planning module (Philips Healthcare, Hamburg, Germany)

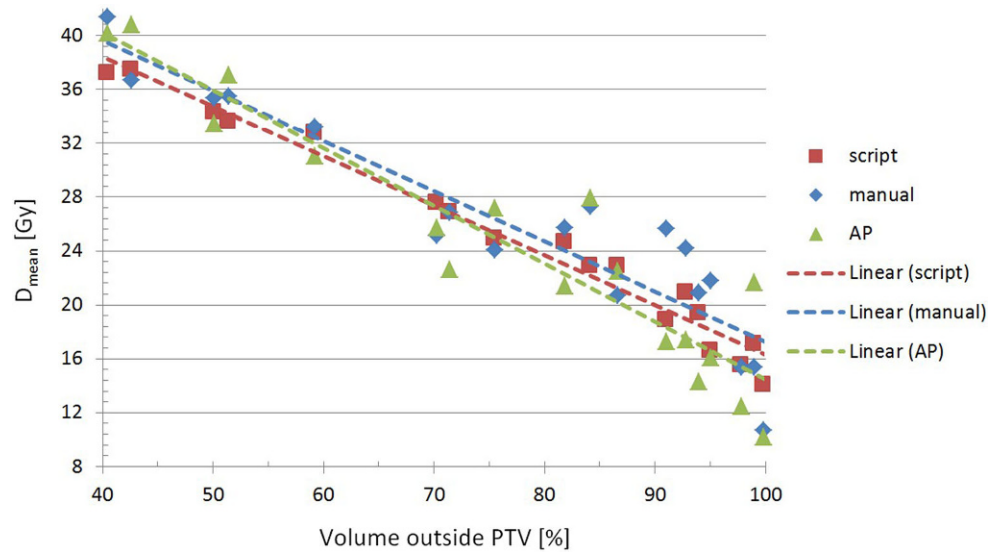
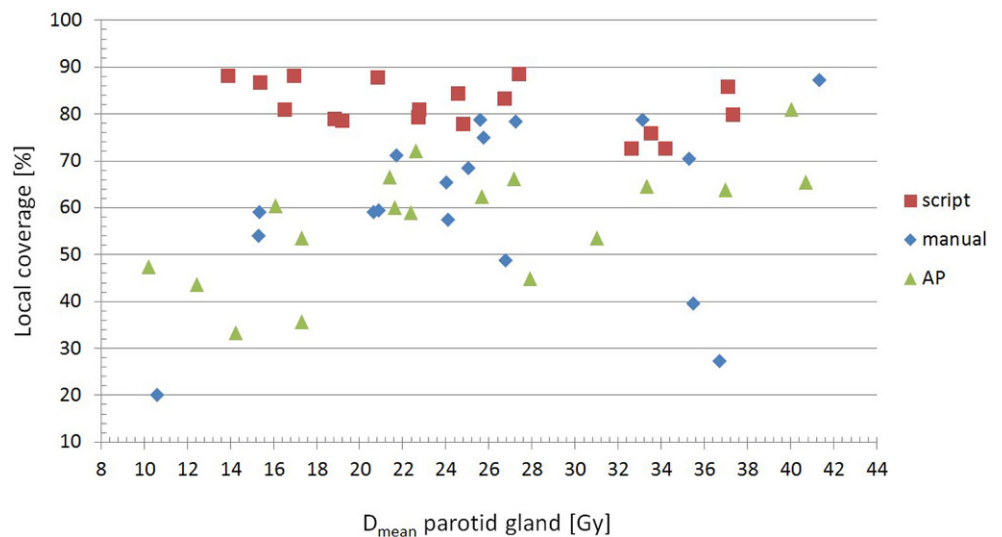


Fig. 3 Relationship between local coverage and mean dose (D_{mean}) to the parotid gland. AP Auto-Planning module (Philips Healthcare, Hamburg, Germany)



can be seen for manual, scripted, and AP plans, but each of them with a different slope or offset:

$$D_{\text{mean}} [\text{Gy}] = 54.517 - 0.3729 \cdot V_{\text{par}} \text{ (manual)}$$

$$D_{\text{mean}} [\text{Gy}] = 53.151 - 0.3687 \cdot V_{\text{par}} \text{ (script)}$$

$$D_{\text{mean}} [\text{Gy}] = 57.441 - 0.43 \cdot V_{\text{par}} \text{ (AP)}$$

where V_{par} is the volume of the parotid gland outside the PTV in % ($40\% < V_{\text{par}} < 100\%$).

It is not possible to decide whether a certain parotid mean dose results from a steeply sloping dose gradient or from simply disregarding PTV coverage in that area. To overcome this limitation, “local parotid coverages” (as defined in Table 1) were also analyzed for each treatment plan.

The evaluated mean doses are 25.84 Gy (manual), 24.79 Gy (script), and 24.37 Gy (AP), where AP has its largest benefits at higher fractions of parotid volume out-

side the PTV (Fig. 2). However, AP has the worst mean local coverage (57.32%). Manual plans are only slightly better (60.80%). The scripted plans seem to be a good compromise between both aspects: a mean local coverage of 81.27% is combined with a mean dose of 24.79 Gy (Fig. 3). When GTV or lymph nodes are located close to the parotid glands, the local coverage should of course be even higher.

Low-priority organs at risk

Without claiming the following list to be complete, larynx, mandible, oral cavity, acoustic area [13], mouth/nose, brain, and skin were classified as low-priority OAR. The data presented in Table 4 correspond to the structures as defined in Table 3. One has to keep in mind that for some patients, data were not available for all OAR (e. g., missing larynx gap;

where “gap” means the rectangular shaped region excluded from PTV intended for sparing the larynx).

Because the mandible and the oral cavity are often completely covered by the PTV, the main goal was to limit the maximum dose to the reference dose. Although the mean maximum doses of all treatment options never exceeded 52.46 Gy in any case, some single outliers with a maximum of up to 54.39 Gy (oral cavity) were observed in manual plans.

Larger differences can be seen for brain, mouth, skin, acoustic area, and larynx. For these OAR, manual plans mostly were inferior to script and AP plans. The mouth (“zone 3”) had best results for AP plans, whereas AP performed poorly for the larynx.

Time requirement

Usually, a scripted plan needs some preliminary work (5–6 min) followed by (up to) six optimization cycles (~30 min each using common settings on the Pinnacle³ SmartEnterprise hardware). No manual intervention was performed for script evaluation and the time needed for generation of a scripted plan therefore summed up to ~3 h. Patient throughput can be increased due to potential parallelization, and the user interaction time is a small fraction of the stated time.

The same can be stated for the AP workflow. The preliminary work takes 5–6 min, accordingly. While Pinnacle is running the AP module, several cycles of IMRT are started internally. Using standard AP settings, it took ~2 h for Pinnacle³ to finish. Of course, AP plans can be parallelized too.

It can be summarized that, for automated plans, the required user input takes less than 5% of the total time required and, therefore, multiple tasks can run at the same time.

For manual plans, one can assume that approximately 1 h of user input distributed throughout the planning process is needed. This will sum up to 4 h (supposing six optimization cycles of 30 min each) until the plan is finished. Despite being error-prone, this workflow is not well suited for parallelization, which thus represents one of the major drawbacks of classic manual planning.

Discussion

The analysis of manual and automated planning techniques for head and neck patients using Pinnacle³ shows that, on average, PTV-related parameters such as coverage, conformity, and homogeneity tend to be better for plans which were generated automatically, as are many other OAR parameters. Some results may be influenced by different op-

timizer settings (e. g., the “compute intermediate dose” option was typically deactivated for manual and AP plans or slightly different maximum iterations per cycle).

AP normally creates acceptable plans which do not need a lot of subsequent interaction; however, some minor problems appear when an OAR lies at least partially inside the PTV and any dose exceeding the target dose should be avoided (mandible, oral cavity). When disabling the “compromise” option, one often encounters doses which are much too low, whereas having the “compromise” hooklet enabled does obviously not affect maximum doses. This can be solved by manual re-optimization. However, when re-planning is required, it may be hard to fix some particular issues. The reason is probably that the generated AP plans are very intensely modulated and are therefore strongly bound to the current solution.

The larynx gap dose was superior for scripted plans. Obviously, sparing this particular region is not easy for Pinnacle³'s Auto-Planning function and manual intervention would be requested in most of the cases.

Because the required rework is often the same, a scripted combination of AP and a final classic IMRT cycle could combine the advantages of both approaches.

Moreover, AP and scripted plans are obviously not able to compete with experienced users when looking at particular plans. Most likely, the knowledge of these planners exceeds the capabilities of any automation. On the other hand, it can be seen that single manual plans did not account for the sparing of some (minor) OAR. Maybe the planners did not know about a certain (minor) OAR constraint at the time when manual plans were treated clinically (2013/14) or simply did not care. This is indicated by a huge standard deviation of almost all manual plan parameters. Scientific conclusions, e. g., of clinical studies, may be complicated when they are based on very heterogeneous plans.

In addition, the parameters listed in Table 3 may not be enough to judge plan quality sufficiently and further criteria should be included for plan evaluation. Particularly, no types of swallowing musculature [14] are yet captured by the scripts. This could be a goal for further work, especially when extending scripts to SIB techniques.

A drawback of scripted plans is the increasing number of MU needed. Strong dose sparing of a larynx gap particularly affects the number of MU. The stronger and more complex the criteria that have to be modulated, the higher the number of MU.

Altogether, a very slight average increase from 358 MU/Gy to 362 MU/Gy accompanied AP plans. Scripted plans required, on average, 486 MU/Gy.

Conclusion

On average, automatically created plans (non-SIB standard head and neck patients generated with Pinnacle³ version 9.10) can be considered to be superior to manually created treatment plans. The score (as defined in Table 3) increased from 59.1 (manual) to 62.3 (AP) and 67.0 points (scripted). Their reproducibility and broad compliance with clinical demands allow even less experienced staff to create plans within an acceptable time, although experienced users are able to excel above these results for particular plans.

Nevertheless, it is important to mention that AP repeatedly failed for one patient. Furthermore, when setting up AP and script templates for the first time, experienced user input is still required, otherwise the benefit of plan automatization may be doubtful.

Among AP and scripted plans there are no relevant differences, and all plans have their own individual benefits for different organs. Pinnacle³'s Auto-Planning algorithm is comparatively easy to run and can also be transferred to other entities without time-consuming instructions. On the other hand, scripts offer more flexibility and may provide better customization.

The mostly minor drawbacks of Pinnacle³'s Auto-Planning algorithm could be covered by applying simple manual corrections or in combination with scripted rework. The combination of scripting and Auto-Planning may provide best results.

In addition, scripts are able to prevent errors due to checking the consistency of manual input. The chance of disregarding OAR is reduced. If needed, plans can be tweaked by experienced physicists, because manual interactions are still possible, albeit not necessary.

It is necessary to mention that these advantages are accompanied by either a very slight or a considerable increase in MU (1% for AP and 36% for scripted plans).

Time consumption can be reduced by a factor of approximately two, or even more when planning is parallelized.

Compliance with ethical guidelines

Conflict of interest S. Speer, A. Klein, L. Kober, A. Weiss, I. Yohannes, and C. Bert declare that they have no competing interests.

Ethical standards This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Thieke C, Küfer KH, Monz M, Scherrer A, Alonso F, Oelfke U, Huber PE, Debus J, Bortfeld T (2007) A new concept for interactive radiotherapy planning with multicriteria optimization: first clinical evaluation. *Radiother Oncol* 85:292–298
2. Buschmann M, Seppenwoolde Y, Wiezorek T, Weibert K, Georg D (2016) Advanced optimization methods for whole pelvic and local prostate external beam therapy. *Phys Med* 32(3):465–473
3. Yang Y, Xing L (2004) Clinical knowledge-based inverse treatment planning. *Phys Med Biol* 49:5101–5117
4. Breedveld S, Storchi PRM, Voet PWJ, Heijmen BJM (2012) iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. *Med Phys* 39:951
5. Quan EM, Chang JY, Liao Z, Xia T, Yuan Z, Liu H, Li X, Wages CA, Mohan R, Zhang X (2012) Automated VMAT treatment planning for stage III lung cancer: How does it compare with IMRT? *Int J Radiat Oncol Biol Phys* 84:e69–e76
6. Wu B, Pang D, Simari P, Taylor R, Sanguineti G, McNutt T (2013) Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: a head-and-neck case study. *Med Phys* 40:021714
7. Geoghegan S (2007) Scripting on the Pinnacle3 treatment planning system. <http://medphysfiles.com/index.php?name=Downloads&get=36>. Accessed 27 June 2016
8. Khaferllari I, Wong E, Bzdusek K, Lock M, Chen JZ (2013) Automated IMRT planning with regional optimization using planning scripts. *J Appl Clin Med Phys* 14:176–191
9. ICRU (2010) Report 83: prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT). *J ICRU* 10(1):NP.3. doi:10.1093/jicru/ndq002
10. Riis B (2013) ICRU83 in Pinnacle. Presentation at the Pinnacle3 user meeting.
11. Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constone LS, Eisbruch A, Bentzen SM, Nam J, Deasy JO (2010) Use of normal tissue complication probability models in the clinic. *Int J Radiat Oncol Biol Phys* 76:S10–S19
12. Wu B, Ricchetti F, Sanguineti G, Kazhdan M, Simari P, Chuang M, Taylor R, Jacques R, McNutt T (2009) Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Med Phys* 36:5497–5505
13. Theunissen EAR, Bosma SCJ, Zuur CL, Spijker R, van der Baan S, Dreschler WA, de Boer JP, Balm AJM, Rasch CRN (2015) Sensorineural hearing loss in patients with head and neck cancer after chemoradiotherapy and radiotherapy: a systematic review of the literature. *Head Neck* 37:281–292
14. Haderlein M, Semrau S, Ott O, Speer S, Bohr C, Fietkau R (2014) Dose-dependent deterioration of swallowing function after induction chemotherapy and definitive chemoradiotherapy for laryngopharyngeal cancer. *Strahlenther Onkol* 190:192–198