Model-based comparison of organ at risk protection between VMAT and robustly optimised IMPT plans

Behrends, C; Haussmann, J; Kramer, P-H; Langendijk, J A; Gottschlag, H; Geismar, D; Budach, W; Timmermann, B

Published in:
Zeitschrift fur medizinische Physik

DOI:
10.1016/j.zemedi.2020.09.003

Model-based comparison of organ at risk protection between VMAT and robustly optimised IMPT plans

Carina Behrends[a,b,c,*], **Jan Haussmann**[d], **P.-H. Kramer**[a,c], **Johannes A. Langendijk**[e], **Holger Gottschlag**[d], **Dirk Geismar**[a,f], **Wilfried Budach**[d], **Beate Timmermann**[a,c,f,g]

[a] West German Proton Therapy Centre Essen (WPE), Essen, Germany
[b] Heinrich-Heine-University, Düsseldorf, Germany
[c] West German Cancer Centre (WTZ), Essen, Germany
[d] Department of Radiation Oncology, Heinrich-Heine-University, Düsseldorf, Germany
[e] Department of Radiation Oncology, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands
[f] Department of Particle Therapy, University Hospital Essen, Essen, Germany
[g] German Cancer Consortium (DKTK), Germany

Eingegangen am 20. Februar 2020; akzeptiert am 18. September 2020

Abstract

The comparison between intensity-modulated proton therapy (IMPT) and volume-modulated arc therapy (VMAT) plans, based on models of normal tissue complication probabilities (NTCP), can support the choice of radiation modality. IMPT irradiation plans for 50 patients with head and neck tumours originally treated with photon therapy have been robustly optimised against density and setup uncertainties. The dose distribution has been calculated with a Monte Carlo (MC) algorithm. The comparison of the plans was based on dose-volume parameters in organs at risk (OARs) and NTCP-calculations for xerostomia, sticky saliva, dysphagia and tube feeding using Langendijk's model-based approach. While the dose distribution in the target volumes is similar, the IMPT plans show better protection of OARs. Therefore, it is not the high dose confirmation that constitutes the advantage of protons, but it is the reduction of the mid-to-low dose levels compared to photons. This work investigates to what extent the advantages of proton radiation are beneficial for the patient's post-therapeutic quality of life (QoL). As a result, approximately one third of the patients examined benefit significantly from proton therapy with regard to possible late side effects. Clinical data is needed to confirm the model-based calculations.

Keywords: Proton therapy, Photon therapy, Head and neck cancer, NTCP-models

1 Introduction

Radiation therapy (RT) is an established therapeutic modality in the treatment of head and neck tumours. A substantial proportion of the patients treated with RT may suffer from radiation-induced late side effects affecting their quality of life (QoL) [1,2]. An important question in medical physics research is to what extent radiation-induced side effects could be avoided. In addition to conventional and, thus, widely used RT with photons, radiotherapy with protons is a promising technology due to the dose deposition over a narrow, well-defined range of depth. In comparison to three dimensional conformal radiation therapy (3D-CRT), state-of-the-art methods of radiation therapy, such as intensity modulated radiation therapy (IMRT) and volumetric modulated arc therapy (VMAT), can reduce dose in organs at risk (OARs), whereby

* Corresponding author: Carina Behrends West German Proton Therapy Centre Essen (WPE), Essen, Germany.
E-mail: carina.behrends@uk-essen.de (C. Behrends).

radiation-induced side effects can be reduced. Both intensity modulated proton therapy (IMPT) and VMAT offer good capabilities for protecting OARs in the treatment of head and neck tumours in comparison to 3D-CRT [3,4]. When comparing proton beams with photon beams, the dosimetric benefit of proton beams lies in their finite range in tissue, while a dosimetric limitation is the broader penumbra compared to photon beams. Proton radiation can potentially deposit lower dose in normal tissue, while the target volume coverage is comparable to photon radiation [5]. In-silico planning comparison studies have confirmed that a reduction in radiation-induced side effects can be achieved with protons [6,7]. The comparison between IMPT and VMAT plans based on models of normal tissue complication probabilities (NTCP), in addition to dose-volume histograms, might support the choice of radiation technique [8]. From a physical point of view, IMPT can yield clinical benefits concerning local tumour control and sparing of normal tissue. This work investigates whether these advantages are beneficial for the patient's post-therapeutic QoL. The relationship between dose distribution in OARs and the development of radiation-induced side effects is described in NTCP-models, with the NTCP-value describing the risk of a given side effect. In general, the NTCP-value will rise with increasing dose to an OAR. Nevertheless, not every dose reduction in the OAR manifests itself in the same reduction of the corresponding NTCP-value (the relationship is a sigmoid curve). In this work, a comparison between IMPT and VMAT is presented with regard to their potential risk reduction for the OAR in the treatment of head and neck tumours. A model-based approach with NTCP-models regarding frequently reported side effects of radiotherapy was used to identify patients benefiting from proton therapy [8].

2 Materials and methods

2.1 Patients and regions of interest

In this in-silico plan comparison study, VMAT radiation plans of clinical patients from the clinic of radiation oncology at the University Hospital Düsseldorf (UKD) were compared with corresponding IMPT plans from the West German Proton Therapy Centre Essen (WPE). These IMPT plans were prepared specifically for this plan comparison, but would have also been clinically acceptable and applicable. The cohort consisted of 50 patients with head and neck tumours of different localisations. Table 1 shows the more detailed localisations of the primary tumour and the respective number of patients. 25 patients underwent adjuvant RT and the other 25 patients primary RT. In addition, in 39 patients the lymphatic vessels were irradiated bilaterally and in 11 patients only unilaterally. Inclusion criteria for the selected cases were patients treated with VMAT for cancers of the head and neck region. The target volumes included treatment of the primary site as well as the uni- and bilateral lymphatic regions. Patients with regional nodal irradiation and patients where the anatomy was insufficiently recognizable for an accurate delineation of the OARs were excluded. In addition, there were no patients in the cohort with disruptive dental artefacts, because of known ineligibility for IMPT.

Table 1
Tumour localisation and corresponding number of patients.

Localisation	Number of patients
Oral cavity	14
Oropharynx	13
Nasopharynx	1
Hypopharynx	8
Larynx	9
Nasal cavities and paranasal sinuses	4
Salivary glands	1

Each case was delineated at the UKD according to local clinical standards. The planning target volumes (PTVs) were created with an isovolumetric 3 mm margin expansion to the clinical target volumes (CTVs). The patients were planned for a simultaneous integrated boost (SIB) technique, whereby the prescribed dose levels listed below correspond to the biological-weighted dose considering an relativ biological effectiveness (RBE) of 1.1 for protons. In the following, all dose values are the RBE weighted dose and short noted as dose.

Three target volumes were identified: a low risk PTV including the elective lymphatic levels with a prescribed dose of 52.8 Gy (PTV52.8), a medium risk PTV including involved lymph levels and the primary gross tumour volume (GTVp) plus a 10 mm anatomically adapted margin with a prescribed dose of 59.4 Gy (PTV59.4). The high risk PTV included the GTVp and the nodal gross tumour volume (GTVn) plus a 5 mm anatomically adapted margin with the prescription of 70 Gy (PTV70) in primary RT cases, or the tumour bed with a 5 mm adapted margin with a prescribed dose of 66 Gy (PTV66). In addition, relevant OARs were contoured and the critical structures of the spinal cord and brain stem had an additional isovolumetric margin of 3 mm for treatment planning optimisation. One half of the patients were treated with 33 fractions of 1.6 Gy in PTV52.8, 1.8 Gy in PTV59.4 and 2.0 Gy in PTV66 up to a total dose of 66 Gy. The other 50% of the patients had the same prescription for PTV52.8 and PTV59.4 but instead of the PTV66, the PTV70 were treated with 35 fractions of 2.0 Gy up to a total dose of 70 Gy. Thus, these patients received a sequential boost with two fractions of 2.0 Gy in the smallest PTV compared to patients with the total dose of 66 Gy.

2.2 Treatment planning

Planning computed tomography (CT) scans were obtained in the UKD with a slice thickness of 3 mm and in supine position. A patient-specific thermoplastic head mask was used for immobilisation. In contrast to the VMAT technique, the IMPT

Table 2
Volumes of interest (VOI) and corresponding planning criteria. These planning criteria based on the RTOG are arranged in decreasing order of priority from top to bottom.

	Volume of interest	Rel. or abs. part of the VOI	Dose [Gy]
Target volume	PTV66	> 95%	62.7
		< 2%	72.6
	PTV70	> 95%	66.5
		< 2%	77.0
	PTV59.4	> 95%	56.4
	PTV52.8	> 95%	50.2
Organ at risk	Spinal cord/spinal cord +3 mm	< 1/3	45.0
		< 2 cm^3	54.0
	Parotid glands	in the mean	< 26
		< 50%	30.0
	PCM superior	in the mean	< 50
	Supraglottic area	in the mean	< 40
	PCM inferior	in the mean	< 30
	Cricopharyngeal muscle	in the mean	< 30
	Submandibular glands	in the mean	< 30
	Soft palate	in the mean	< 40
	Oesophageal inlet muscle	in the mean	< 30
	Oral cavity	in the mean	< 37
	Sublingual glands	in the mean	< 30
	Glottic larynx	in the mean	< 40
	PCM medius	in the mean	< 30
	Brain stem/brain stem +3 mm	< 1/3	54.0
		< 2 cm^3	60.0
	Optic nerve left/right	< 1 cm^3	54.0
	Chiasm	< 1 cm^3	54.0
	Eye left/right	< 1 cm^3	50.0
	Macula left/right	< 1 cm^3	45.0
	Lens left/right	< 1 cm^3	12.0
	Inner ear left/right	< 1 cm^3	50.0
	Mandibula	< 1 cm^3	70.0

technique requires the body-shaped storage aid BosFrame (company Qfix) due to the implementation of the treatment. For this reason, the table used in the planning CT carried out at the UKD had to be changed when the IMPT comparison plans were created. However, the change in table material had no dosimetrical effect on the comparison plan but was necessary as the edges of the treatment table would have collided with the snout. Furthermore, dental structures with Hounsfield unit (HU) values greater than or equal to those of aluminum were overwritten with aluminum and artifacts were overwritten with water, as this is clinical practice at WPE. The optimisation specifications and dose concepts were identical for both techniques in advance, based on local standards and the Radiation Therapy Oncology Group (RTOG) protocol (see Table 2) [9]: at least 95% of the planning target volumes should be covered by 95% of the prescribed dose and the total dose was 66 Gy in 33 fractions or 70 Gy in 35 fractions. In addition, planning target volumes should receive their prescribed dose level on average. Table 2 lists the planning criteria in order of priority. While the target coverage had highest priority, doses to OARs were reduced as far as the target coverage of $V_{95\%} > 95\%$ was not compromised.

2.2.1 VMAT

All clinical VMAT plans were re-optimised for the relevant target volumes and OARs in order to allow a planning comparison, using the same constraints as given in the IMPT treatment planning, which are listed in Table 2. Dedicated swallowing-sparing IMRT was not the standard of care during the inclusion time and thus not all relevant OARs were included in the optimisation process. All re-optimised plans were assessed by a physician and deemed clinically acceptable.

The VMAT irradiation plans were calculated using the Eclipse planning system (version 13.6, Varian medical systems, Palo Alto, California, USA) at UKD. They were created using two partial or full arcs, depending on the tumour location and optimised for a Varian TrueBeam STx machine. The Acuros algorithm was used for dose calculation. Moreover, the optimisation process was performed on a modified CT where air cavities in close proximity to the target volumes were substituted with water equivalent density. Evaluation PTVs were created with a 2 mm margin from the skin (except for treatments with a bolus) in order to maintain a dose

build-up. The dosimetric goals for target coverage were according to ICRU 83 [10].

2.2.2 IMPT

The corresponding IMPT treatment plans were optimised in the RayStation planning system (version 7, RaySearch Laboratories, Stockholm, Sweden) at WPE, with local clinical standards [11]. They consisted of five beams with gantry angles 0°, 70°, 120°, 240° and 290°. Depending on the tumour growth area, sometimes only three beams and slightly different gantry angles were used. Concerning the spot pattern, the spot spacing was automatically calculated with the radial spread in the Bragg peak for a specific energy. This has lead to the clinically used range of about 0.7 cm to 1.2 cm. The energy layer spacing was automatically calculated depending on the Bragg peak width. Minimum weights per spot were around 0.027 monitor units/fraction. Since the minimum available proton energy was 100 MeV, a range shifter (material: polymethyl methacrylate) with a water equivalent thickness of 7.4 cm was used to reduce the range of the proton beams. Due to the different storage aids at UKD and WPE, the IMPT planning was created without treatment tables. The RBE weighted dose was calculated with a Monte Carlo (MC) algorithm (MC dose calculation engine: v4.1) [12]: 50000 ions per spot were considered with the mean relative statistical uncertainty (MRSU) reaching 0.5% and a calculation grid of $2 \cdot 2 \cdot 2\,\text{mm}^3$. The plans should be robust against uncertainties in the range and patient position [13,14]. Random perturbation analyses were performed to check the robustness of the IMPT plans. Values of ±3.5% deviation in density and ±3 mm setup uncertainty were selected as favourable perturbation scenarios for the head and neck area [15]. The plan was sufficiently robust if the worst-case scenario still met the clinical objective. Ideally, 100% of the CTV was irradiated with 95% of the prescribed dose. It was desired that the prescribed dose still covered more than 98% of the CTV for perturbation. This corresponded to the near minimum dose $D_{98\%}$ defined by the ICRU.

In order to be able to compare the IMPT plans with the VMAT plans, the PTVs of the IMPT plans were evaluated, although robust optimisation was performed for the CTV.

2.3 Statistical analysis of the average OAR dose

In order to compare the dose delivered to OARs for VMAT and IMPT, a hypothesis test, in this case the *t*-test for paired samples, was performed. This test checked whether VMAT generated a higher dose in the respective OAR than IMPT. The null hypothesis was that there was no difference in mean OAR dose between VMAT and IMPT: $\Delta\mu = \mu_{\text{VMAT}} - \mu_{\text{IMPT}} = 0\,\text{Gy}$. The level of significance was initially set to 5% ($\alpha = 0.05$). This test was performed separately for the 66 Gy and 70 Gy dose prescriptions, each with a sample of 25 patients and 14 OARs.

2.4 NTCP-calculation

Langendijk et al. presented the model-based approach as a method for the selection of patients for proton therapy with the primary goal to reduce the side effects [8]. To assess the radiation-induced late side effects limiting the patient's QoL, NTCPs of the most frequent late side effects of radiological treatments in the head and neck region were calculated for both treatment techniques. These endpoints include moderate to severe xerostomia six months after RT, sticky saliva at the end of RT, grade 2–4 dysphagia according to the RTOG/European Organisation for Research and Treatment of Cancer (EORTC) six months after RT and tube feeding dependence six months after RT [16–18].

The dose–volume relationships in the target volumes and OARs of the two irradiation techniques were compared. Based on this, the mean dose deposited in the OAR was evaluated. However, since a dose reduction can have different effects on the patient, NTCP-calculations are decisive. The following NTCP-models were used, which are a measure of the relationship between dose distribution in the risk structures and the development of radiation-induced side effects. Generally, the NTCP-value is calculated corresponding to [16]:

$$\text{NTCP} = (1 + e^{-S})^{-1},$$

where

$$S = \beta_0 + \sum_{i=1}^{n} \beta_i \cdot x_i.$$

According to the side effect, the exponent S consists of regression coefficients β and variables x_i. For the dosimetric variables the dose in units of Gy can be filled in, while the non-dosimetric variables are either 0 (= no) or 1 (= yes). For the individual patients the NTCP-value for xerostomia and sticky saliva can be calculated using the formula for the according exponents [16]:

$$\begin{aligned} S_{\text{xero}} &= -1.443 + (\overline{D}_{\text{contralat. parotid}} \cdot 0.047) \\ &\quad + (\alpha \cdot 0.720) \end{aligned} \tag{1}$$

$$\begin{aligned} S_{\text{sticky}} &= -3.243 + (\overline{D}_{\text{contralat. submandibul.}} \cdot 0.075) \\ &\quad + (\overline{D}_{\text{sublingual gl.}} \cdot (-0.060)) \\ &\quad + (\overline{D}_{\text{soft palate}} \cdot 0.026). \end{aligned} \tag{2}$$

The mean dose of the corresponding OARs is marked with $\overline{D}$, α is the reference value for xerostomia [16].

According to [17], the NTCP-value for the physician-rated swallowing dysfunction can be calculated with the formula

$$\begin{aligned} S_{\text{dys}} &= -6.09 + (\overline{D}_{\text{PCM superior}} \cdot 0.057) \\ &\quad + (\overline{D}_{\text{supraglot. larynx}} \cdot 0.037). \end{aligned} \tag{3}$$

The exponent for the NTCP-calculation for tube feeding is determined by [18]:

$$\begin{aligned} S_{\text{tube}} &= -11.70 + (T \cdot 0.43) \\ &\quad +(W_{\text{moderate}} \cdot 0.95) + (W_{\text{severe}} \cdot 1.63) \\ &\quad +(R \cdot 1.20) + (C \cdot 1.91) + (R_{\text{cet}} \cdot 0.56) \\ &\quad +(\overline{D}_{\text{PCM superior}} \cdot 0.071) \\ &\quad +(\overline{D}_{\text{PCM inferior}} \cdot 0.034) \\ &\quad +(\overline{D}_{\text{contralat. parotid}} \cdot 0.006) \\ &\quad +(\overline{D}_{\text{cricopharyn. muscle}} \cdot 0.023). \end{aligned} \tag{4}$$

Here, T is the advanced tumour stage, which is 0 for tumour stage 1 or 2 and is 1 for stage 3 or 4. The non-dosimetric variable W_{moderate} is a moderate weight loss (1–10% weight loss), W_{severe} a severe weight loss (>10% weight loss), R the accelerated radiotherapy, C the chemo radiation and R_{cet} the radiotherapy plus cetuximab.

For the difference $\Delta\text{NTCP}_{\text{endpoint}} = \text{NTCP}_{\text{endpoint,VMAT}} - \text{NTCP}_{\text{endpoint,IMPT}}$, medically meaningful limits from the Dutch national indication protocol for head and neck tumours (Landelijk Indicatie Protocol Protonentherapie (LIPPv2.2)) were chosen that identify patients who benefit from proton therapy. In this study, proton therapy was chosen when one of the three ΔNTCP-values of xerostomia, sticky saliva or dysphagia exceeded the 10% limit or two of them exceeded the 7.5% limit or all of these three values added up to 15%. For $\Delta\text{NTCP}_{\text{tube}}$ the limit for proton therapy was 5%. With this model, the decision was made whether to use proton or photon therapy for the 50 patients on the basis of the calculated ΔNTCP. The sum of all NTCP-calculations $\Delta\text{NTCP}_{\text{tot}}$ was defined as

$$\begin{aligned} \Delta\text{NTCP}_{\text{tot}} &= \sum_{\text{all}} \Delta\text{NTCP}_{\text{endpoint}} \\ &= \Delta\text{NTCP}_{\text{xero}} + \Delta\text{NTCP}_{\text{sticky}} \\ &\quad +\Delta\text{NTCP}_{\text{dys}} + \Delta\text{NTCP}_{\text{tube}}. \end{aligned} \tag{5}$$

3 Results

3.1 Dose-statistics

As shown in Figure 1, differences between the simulated dose distributions of VMAT and IMPT plans are discernible, although their relevance is initially open. These differences are also apparent in the applied doses to the OAR (see Figure 2). The high dose distribution differs slightly or not at all, while in mid-to-low dose areas the differences are more pronounced. The dose coverage of the target volumes is approximately the same for protons and photons (see Figure 3). The following statistical analysis tests and confirms these assumptions.

Figure 1 displays the dose distributions of the IMPT and the VMAT exemplarily in the sagittal and transverse sectional plane. To measure dose conformity, the conformity index (CI) is calculated for different dose levels. Here, $\Delta\text{CI} = \text{CI}_{\text{IMPT}} - \text{CI}_{\text{VMAT}}$ indicates the difference in conformity index of IMPT and VMAT. ΔCI(PTV70) for the 66.5 Gy isodose is 0.13, for the 56.4 Gy isodose and the 50.2 Gy isodose ΔCI(PTV59.4) and ΔCI(PTV52.8) are both 0.10, but for the blue 35 Gy isodose ΔCI(PTV52.8) is 0.16. Therefore, the high dose distribution of the IMPT and VMAT plan adds up to nearly the same conformity, while the IMPT plan shows better mid-to-low dose conformity and better OAR sparing, as can be seen in the submandibular glands, for example. The distribution of the mean dose in the critical structures is shown in Figure 2. The values of the dose statistics scatter very widely, especially for the organs in the oral cavity as well as the swallowing structures. According to the planning criteria in Table 2, after sufficient target volume coverage, the best possible OAR sparing and compliance with the defined OAR planning criteria was also ensured. However, there are some cases, both for VMAT and IMPT, where the mean dose is higher than the limit value. In general, the median value of the applied dose in the individual OARs is lower for IMPT than for VMAT, except for the contralateral parotid gland. Furthermore, $D_{2\%}$ is applied for the spinal cord in addition to the average dose. While the mean dose is significantly lower in the IMPT compared to the VMAT, the techniques differ only slightly when comparing the $D_{2\%}$: In the 66 Gy total dose plans, the median near-maximum dose in the spinal cord of the IMPT is about 2.5 Gy higher than in the VMAT, while in the 70 Gy plans the median of the IMPT is about 1 Gy lower than the VMAT. The integral dose calculated by mean body dose times body volume shows a difference between the VMAT and the IMPT: It is 474.9 Gy · L $\left(\min = 52.29\,\text{Gy}\cdot\text{L};\ \max = 1.584\cdot10^3\,\text{Gy}\cdot\text{L}\right)$ for IMPT and 525.6 Gy · L $\left(53.11\,\text{Gy}\cdot\text{L};\ 1.585\cdot10^3\,\text{Gy}\cdot\text{L}\right)$ for VMAT, whereby the minimum and maximum values are additionally indicated. This means that the integral dose is on average almost 10% lower for IMPT than for VMAT.

The t-test for paired samples rejects the null hypothesis ($\Delta\mu = 0$ Gy) in both cases (total dose 66 Gy and 70 Gy) at a significance level of 5%. The p-value for the total dose of 66 Gy is $4.78\cdot10^{-6}$, and for 70 Gy the p-value is $1.39\cdot10^{-4}$. Therefore, the null hypothesis is actually rejected at a significance level of 0.1%. The 95% confidence interval for $\Delta\mu$ is (1.98 Gy; 7.16 Gy) for 66 Gy and (2.74 Gy; 6.49 Gy) for 70 Gy. Consequently, this is a highly significant result and it confirms the assumption of better OAR protection using IMPT made in Figure 1. To what extent these dose reductions are relevant for the patient's QoL is shown in the following NTCP-calculations (Section 3.2).

The mean dose distribution and $D_{95\%}$ coverage of the target volumes for all patients is demonstrated in Figure 3. The patients with a total dose of 66 Gy and those with a total dose

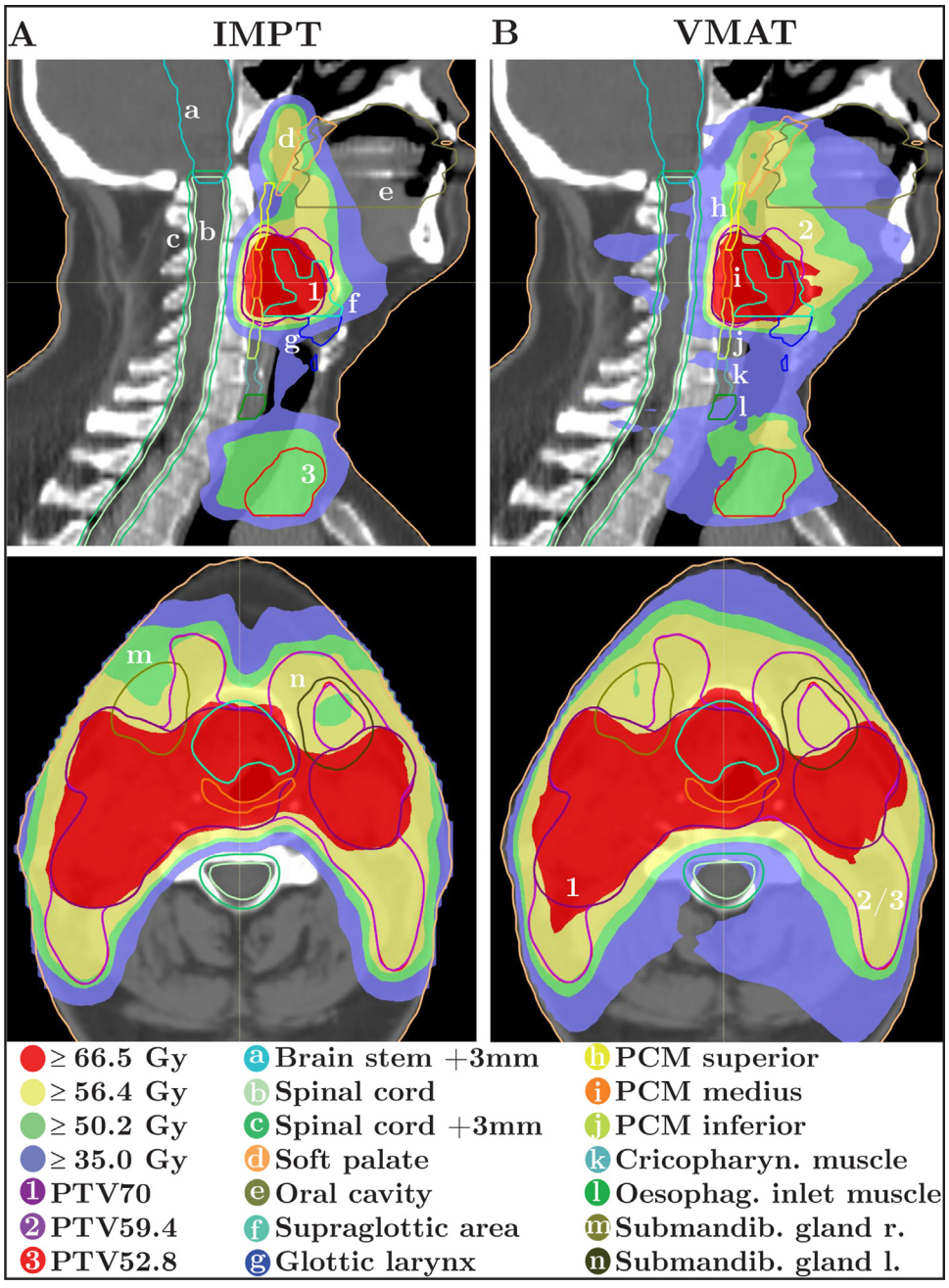

Figure 1. Simulated dose distribution from the planning system RayStation, example for one patient. **A** illustrates the dose distribution from the IMPT plan (RBE weighted) and **B** shows the dose distributions from the VMAT plan. The upper images are in the sagittal view, while the lower images show the distribution in a transversal plane. The isodoses $D_{95\%}$ of the dose levels of the SIB are marked in the colours red, yellow and green. The blue isodose indicates the 50% dose level of the prescription dose.

of 70 Gy were analysed separately. Since the target volumes of the SIB plans lie within each other, the subtracted volumes for PTV52.8 and PTV59.4 to the volume of the next higher dose level, hereafter referred to as ΔPTV52.8 and ΔPTV59.4, were evaluated accordingly. For most cases, the mean dose is slightly above the dose level of the target volume, whereas the target coverage is within the predefined constraints for $D_{95\%}$ (see Table 2). The crosses in Figure 3 correspond to patients in whom an OAR conflicts with the target volume coverage and a compromise had to be reached despite the planning criteria. These statistics confirm the almost identical tumour coverage of both irradiation techniques in the high-dose range.

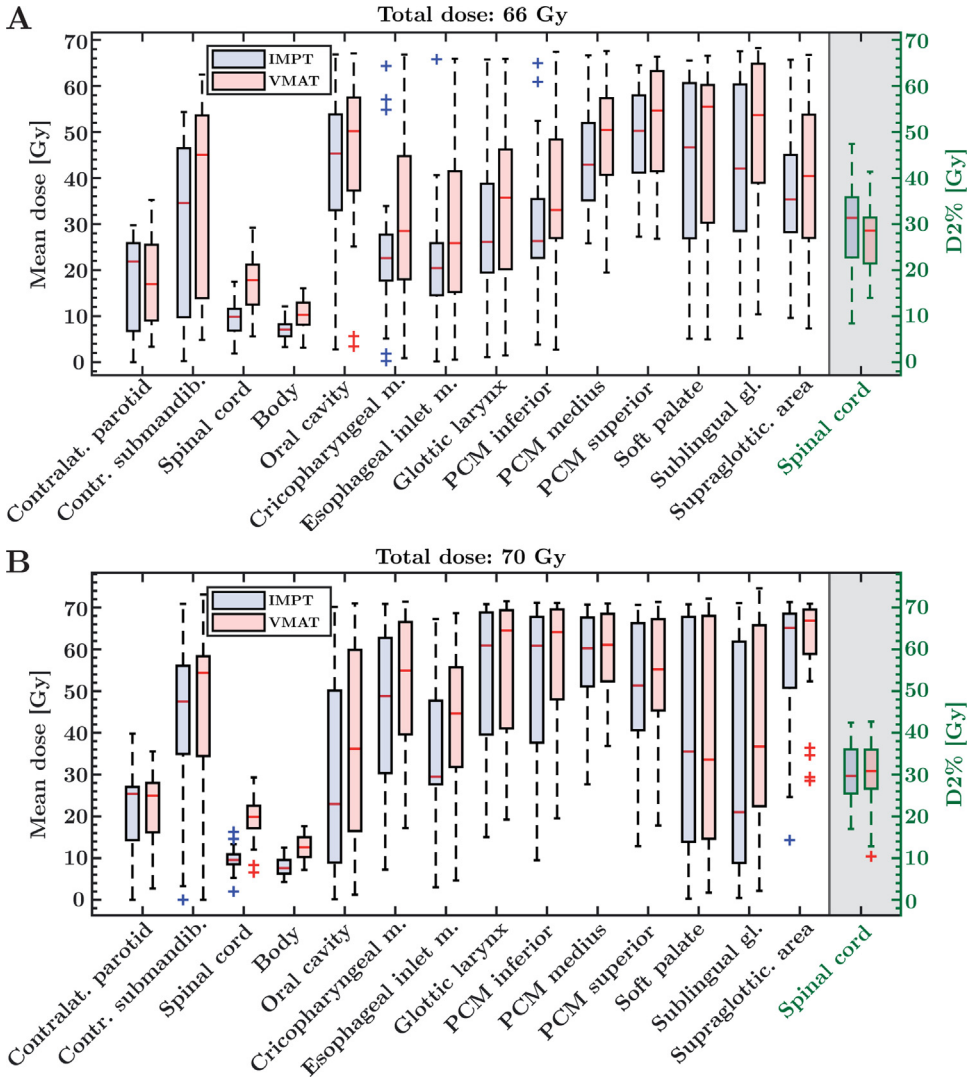

Figure 2. Mean dose distribution of all organs at risk and over all patients, with the prescribed total dose of 66 Gy (**A**) and 70 Gy (**B**). In addition, $D_{2\%}$ is shown for the serial organ spinal cord in green (right *y*-axis). On the boxes, the central red marker indicates the median, and the lower and upper edges of the box indicate the 25th and 75th percentile respectively. The plotted whisker extends to 1.5 times the length of the box and the adjacent value, which is the most extreme data value that is not an outlier (marked with crosses).

3.2 NTCP-statistics

The OAR dose reductions obtained with IMPT show numerically reduced NTCP-values for RTOG grade 2–4 swallowing dysfunction and tube feeding (Figure 4). The median NTCP of dysphagia could be reduced from 24% to 18% and the value of tube feeding from 6% to 3%. In contrast, the results of the NTCP-calculation of xerostomia and sticky saliva show an advantage in photon therapy. For some patients, the value of ΔNTCP was well above the marked lines, which indicate the established limits for the decision of therapy.

Based on the resulting calculations of the individual ΔNTCP, 16 of 50 patients treated with photons could have been advised for proton therapy because of a possible clinical benefit (Figure 4) if decisions were taken on the basis of the limit values introduced. Figure 5 shows the $\Delta NTCP_{tot}$-value for each patient, which is based on Eq. (5) giving the sum over all individual ΔNTCP-values of this model-based approach. For 10% of the patients – patient 2, 12, 14, 21 and 42 – the advantage of proton therapy would have been particularly high ($\Delta NTCP_{tot} > 30\%$).

4 Discussion

This comparative planning study demonstrated that the dose distributions in the OARs for IMPT are significantly lower

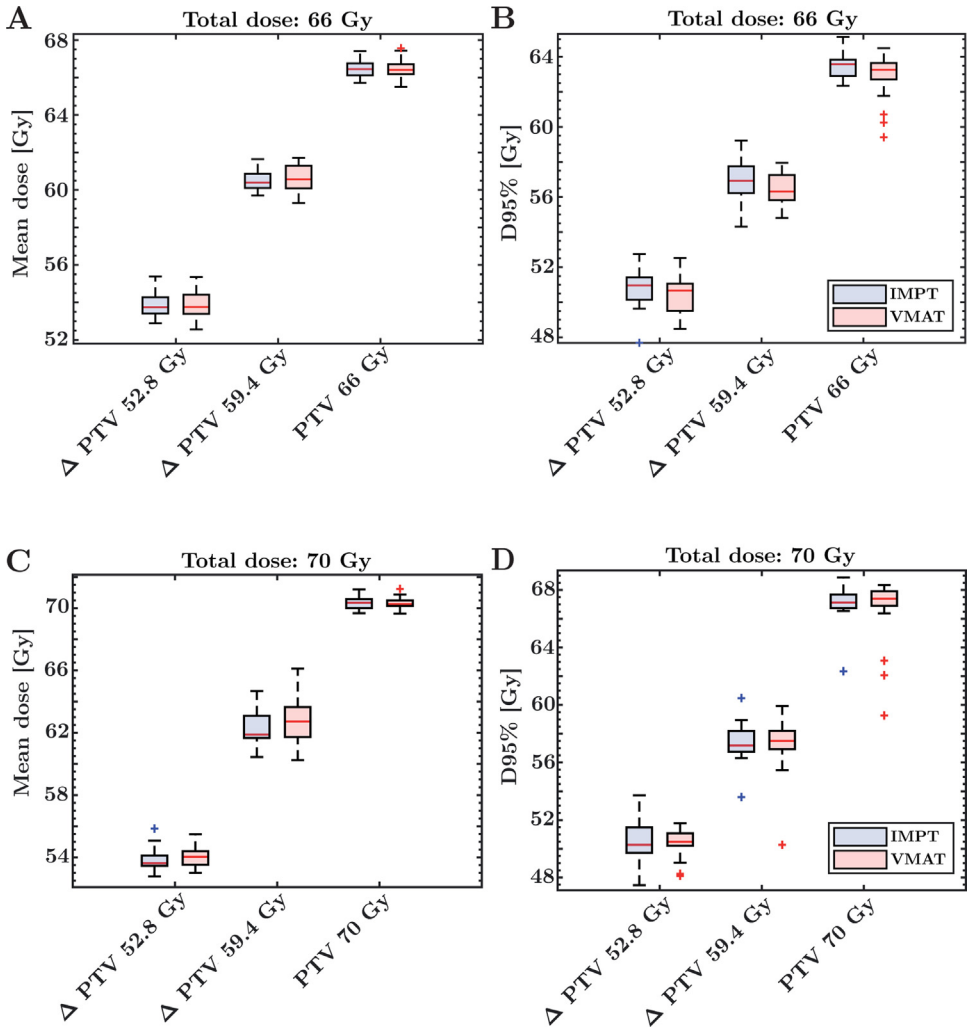

Figure 3. Dose distribution of the target volumes ΔPTV52.8, ΔPTV59.4 and PTV66/PTV70 for mean dose and $D_{95\%}$ evaluated over all patients. Since the target volumes are located in each other, the subtracted volumes ΔPTV52.8 and ΔPTV59.4 were used to clarify the target coverage. **A** shows the mean dose und **B** the $D_{95\%}$ in each case for the total dose of 66 Gy. Accordingly, the prescribed total dose of 70 Gy is shown in **C** (mean dose) and **D** ($D_{95\%}$).

than for VMAT, while the PTV dose coverage is the same for both techniques. Based on the NTCP-model analysis, 32% of the patients in the cohort would benefit significantly from proton therapy.

In almost all OARs the advantage of the finite range in tissue of the protons was shown. Especially the organs in the mouth and throat area that are often close to the tumour, but also the spinal cord and the whole body, could be spared better during treatment with protons. For the contralateral parotid glands, however, the median values of the OAR dose are lower with the VMAT technique. Accordingly, only in the latter case more patients receive a higher dose with IMPT than with VMAT (Figure 2). This fact could be explained by the discrete beam directions of the IMPT. In order to protect the mouth, the lateral fields loaded the parotid glands, while the VMAT uses all coplanar beam directions. For the most part, the OAR dose could be somewhat reduced by the IMPT technique, while the dose coverage in the target volume was about the same for IMPT and VMAT (Figure 3). In any case, the advantage of proton radiation does not lie in the target volume coverage. If the overall dose distribution is compared, the advantage lies clearly in the mid-to-low dose areas in comparison to photons. Other groups have already shown that there may be clinical benefits for these significantly reduced volumes with mid-to-low doses: Miralbell et al. presented that lower integral doses reduce formation of secondary tumours [19]. According to Durante et al. and Bijl et al., the pre-clinical effects can also be affected by a reduced normal tissue dose [20,21].

The NTCP-calculations confirmed the clinical advantage of proton therapy due to the significantly reduced volumes of mid-to-low doses, especially for dysphagia and tube feeding in studied cases. According to the model, one third of the

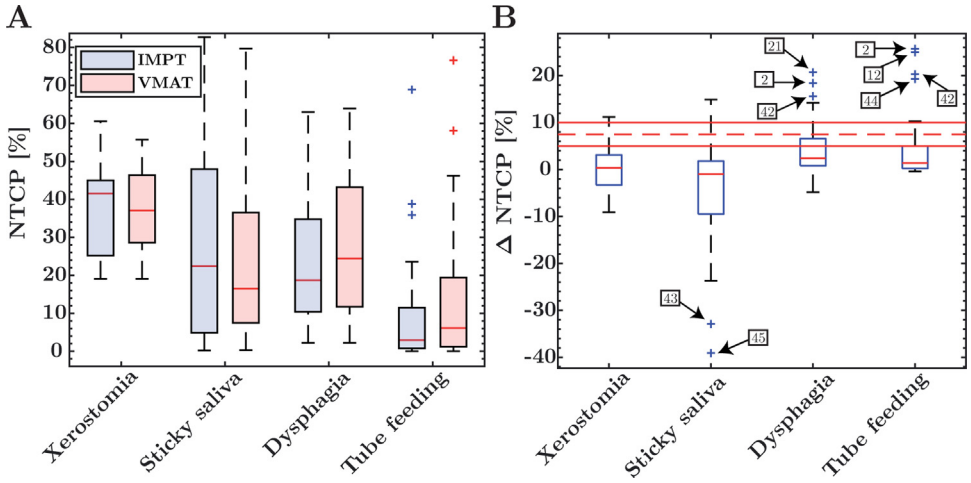

Figure 4. NTCP-statistics concerning xerostomia, sticky saliva, dysphagia and tube feeding over all patients. In **A** the calculated NTCP-values are shown and in **B** the corresponding differences $\Delta\text{NTCP} = \text{NTCP}_{\text{VMAT}} - \text{NTCP}_{\text{IMPT}}$. The numbers characterise the respective patients and the red lines show the limits.

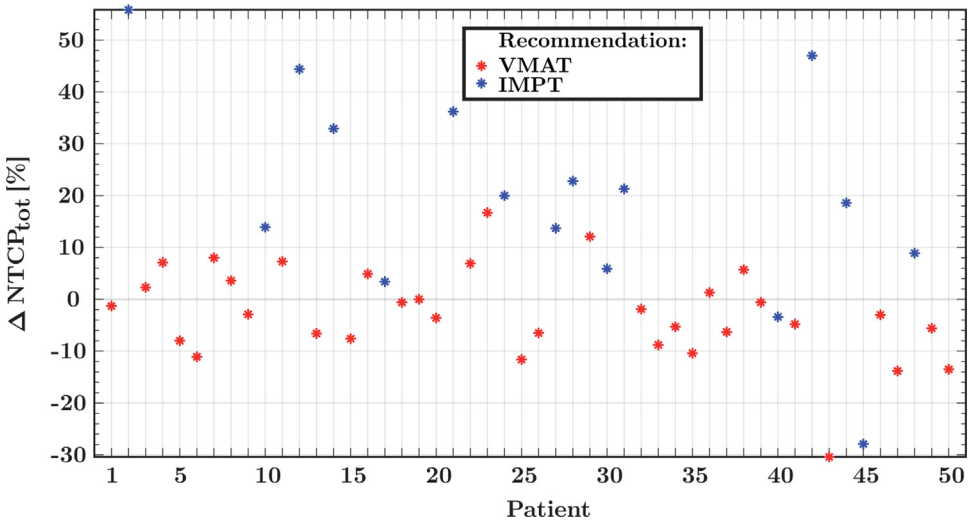

Figure 5. Results of the model-based approach of the $\Delta\text{NTCP}_{\text{tot}}$-value for each patient, calculated by Eq. (5). The colouring displays the individual model based proposal for VMAT (red) and IMPT (blue). This proposal results from the respective $\Delta\text{NTCP}_{\text{endpoint}}$-values and the limit values used in the model (see Section 2.4): An IMPT is adviced when one of the respective $\Delta\text{NTCP}_{\text{endpoint}}$ for xerostomia, sticky saliva or dysphagia exceeded the 10% limit or two of them exceeded the 7.5% limit or all of these three values added up to 15% or the NTCP for tube feeding exceeded the 5% limit.

examined patients would benefit significantly from treatment with protons. Since the decision on the form of therapy was based on the chosen limit values, this choice was a decisive factor.

Due to the limited treatment machines and the higher cost of IMPT, it is essential to identify a subgroup of patients that benefit the most from proton therapy. A pronounced characteristic such as tumour localisation, which predicts the benefit of IMPT, could not be found in the five patients with a significant benefit of proton therapy (patient 2, 12, 14, 21 and 42). The tumours of these five patients, which were well lateralised, were nevertheless irradiated on both sides due to their geometry. Therefore, the dose was deposited in the entire volume during treatment with photons. The protons optimally avoided the internal OARs. The plan comparison therefore suggests that tumours of this type are better suited for proton therapy, but this is only a hypothesis that must be confirmed by further investigations.

Possible errors during evaluation have to be discussed within the framework of the plan comparison. On the one hand, these may have occurred due to the fact that the planning CT had to be modified due to different technical implemen-

tations of the irradiation and different planning systems. On the other hand, an extremely important and critical point is the comparison based on the PTVs, which was defined as the uniform extension of the CTV. This construction of the PTV for proton therapy was not sufficient because range and setup uncertainties may vary depending on the beam direction. Therefore, the IMPT plans were planned robustly on the CTV. For this reason, the PTV evaluation for proton therapy was not entirely correct. In order to be able to compare the plans appropriately, this uncertainty was accepted. In the future, this uncertainty could be corrected by implementing the PTV-less photon and proton treatment planning [22].

The proportion of patients who would be selected to proton therapy in point of fact might be lower than the approximately 1/3 presented in this study because patients with metal fillings or inlays were excluded from the cohort. Furthermore, the model does not include all risks and only considers certain organ doses. For example, the integral dose, which was additionally evaluated in this work, is missing. Besides, the NTCP-models used in the model-based approach have been developed using IMRT data. Therefore, the fundamental question is whether the NTCP-models are transferable to protons.

The variable RBE optimisation was not taken into account in the current work. Therefore, the dose calculations may have led in favour of proton plans when comparing the treatment plans [23]. It is known that the RBE increases with increasing LET and can achieve a value of 1.7 or even 4–6 in the distal fall-off of the SOBP at 2 Gy/fraction [24–26]. Nevertheless, it is difficult to estimate the influence for this plan comparison, since the RBE depends on various factors (dose, LET, cell type, endpoint). Furthermore, the increases in RBE confirmed so far are based on in vitro data and it is still unclear to what extent clinical data confirm cell line data [27]. For subsequent plan comparisons of IMRT and IMPT it would be interesting to consider the RBE weighted dose of protons.

This NTCP-model-based in-silico plan comparison demonstrated the IMPT's potential to reduce normal tissue complications relative to VMAT. However, all results of this work relate to the models used and must therefore be considered with the same reservation as the models themselves. A next step would be to check the photon-based model for applicability and validity with protons and possibly extend it. This has already been investigated for head and neck cancer and brain tumour side effects and rectum morbidity [28–30]. Furthermore, these NTCP-models were based on IMRT, therefore the applicability for VMAT must also be checked. Clinical validation studies are required to confirm whether the calculated dose reductions of the considered OARs result in the reduction of the severity of the late side effects. Based on the results of this work, a follow-up clinical study is planned to further investigate the comparison of proton and photon RT and to validate the NTCP-models for VMAT and IMPT.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The authors gratefully acknowledge the provided language support of E. Den Boer. Moreover, we are indebted to B. Tamaskovics for proof reading from the medical side and to C. Bäumer and J. Wulff for help for discussions.

References

[1] Langendijk JA, Doornaert P, de Leeuw IMV, Leemans CR, Aaronson NK, Slotman BJ. Impact of late treatment-related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. J Clin Oncol 2008;26(22):3770–6, http://dx.doi.org/10.1200/jco.2007.14.6647.

[2] Jensen K, Jensen AB, Grau C. The relationship between observer-based toxicity scoring and patient assessed symptom severity after treatment for head and neck cancer. A Correlative cross sectional study of the DAHANCA toxicity scoring system and the EORTC quality of life questionnaires. Radiother Oncol 2006;78(3):298–305, http://dx.doi.org/10.1016/j.radonc.2006.02.005.

[3] Ghosh G, Tallari R, Malviya A. Toxicity profile of IMRT vs. 3D-CRT in head and neck cancer: a retrospective study. J Clin Diagn Res 2016;10(9), http://dx.doi.org/10.7860/jcdr/2016/21457.8583.

[4] Langendijk JA, Boersma LJ, Rasch CR, van Vulpen M, Reitsma JB, van der Schaaf A, et al. Clinical trial strategies to compare protons with photons. Semin Radiat Oncol 2018;28(2):79–87, http://dx.doi.org/10.1016/j.semradonc.2017.11.008.

[5] van de Water TA, Bijl HP, Schilstra C, Pijls-Johannesma M, Langendijk JA. The potential benefit of radiotherapy with protons in head and neck cancer with respect to normal tissue sparing: a systematic review of literature. Oncologist 2011;16(3):366–77, http://dx.doi.org/10.1634/theoncologist.2010-0171.

[6] Stuschke M, Kaiser A, Pöttgen C, Lübcke W, Farr J. Potentials of robust intensity modulated scanning proton plans for locally advanced lung cancer in comparison to intensity modulated photon plans. Radiother Oncol 2012;104(1):45–51, http://dx.doi.org/10.1016/j.radonc.2012.03.017.

[7] Schwarz M, Pierelli A, Fiorino C, Fellin F, Cattaneo GM, Cozzarini C, et al. Helical tomotherapy and intensity modulated proton therapy in the treatment of early stage prostate cancer: a treatment planning comparison. Radiother Oncol 2011;98(1):74–80, http://dx.doi.org/10.1016/j.radonc.2010.10.027.

[8] Langendijk JA, Lambin P, Ruysscher DD, Widder J, Bos M, Verheij M. Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach. Radiother Oncol 2013;107(3):267–73, http://dx.doi.org/10.1016/j.radonc.2013.05.007.

[9] Machtay M, Thorstad W, Nquyen-Tân PF, Siu LL, Holsinger FCF, Quon H, et al. A phase III study of postoperative radiation therapie imrt +/- cetuximab for locally-advanced resected head and neck cancer, radiation therapy oncology group RTOG 0920; 2015.

[10] International I. Commission on radiation units, measurements, prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT). J ICRU 2010;10(1), http://dx.doi.org/10.1093/jicru/ndq002. NP.3-NP.

[11] Bäumer C, Geismar D, Koska B, Kramer PH, Lambert J, Lemke M, et al. Comprehensive clinical commissioning and validation of the RayStation treatment planning system for proton therapy with active

scanning and passive treatment techniques. Phys Med 2017;43:15–24, http://dx.doi.org/10.1016/j.ejmp.2017.09.136.
[12] Saini J, Maes D, Egan A, Bowen SR, James SS, Janson M, et al. Dosimetric evaluation of a commercial proton spot scanning Monte-Carlo dose algorithm: comparisons against measurements and simulations. Phys Med Biol 2017;62(19):7659–81, http://dx.doi.org/10.1088/1361-6560/aa82a5.
[13] Fredriksson A. A characterization of robust radiation therapy treatment planning methods-from expected value to worst case optimization. Med Phys 2012;39(8):5169–81, http://dx.doi.org/10.1118/1.4737113.
[14] Fredriksson A, Forsgren A, Hårdemark B. Maximizing the probability of satisfying the clinical goals in radiation therapy treatment planning under setup uncertainty. Med Phys 2015;42(7):3992–9, http://dx.doi.org/10.1118/1.4921998.
[15] Liu W, Frank SJ, Li X, Li Y, Park PC, Dong L, et al. Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers. Med Phys 2013;40(5):051711, http://dx.doi.org/10.1118/1.4801899.
[16] Beetz I, Schilstra C, van den Schaaf A, van den Heuvel ER, Doornaert P, van Luijk P, et al. NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors. Radiother Oncol 2012;105(1):101–6, http://dx.doi.org/10.1016/j.radonc.2012.03.004.
[17] Christianen ME, 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, http://dx.doi.org/10.1016/j.radonc.2011.08.009.
[18] Wopken K, Bijl HP, van der Schaaf A, van der Laan HP, Chouvalova O, Steenbakkers RJ, et al. Development of a multivariable normal tissue complication probability (NTCP) model for tube feeding dependence after curative radiotherapy/chemo-radiotherapy in head and neck cancer. Radiother Oncol 2014;113(1):95–101, http://dx.doi.org/10.1016/j.radonc.2014.09.013.
[19] Miralbell R, Lomax A, Cella L, Schneider U. Potential reduction of the incidence of radiation-induced second cancers by using proton beams in the treatment of pediatric tumors. Int J Radiat Oncol Biol Phys 2002;54(3):824–9, http://dx.doi.org/10.1016/s0360-3016(02)02982-6.
[20] Durante M, Orecchia R, Loeffler JS. Charged-particle therapy in cancer: clinical uses and future perspectives. Nat Rev Clin Oncol 2017;14(8):483–95, http://dx.doi.org/10.1038/nrclinonc.2017.30.
[21] Bijl HP, Luijk Pv, Coppes RP, Schippers JM, Konings AW, Der Kogel AJv. Regional differences in radiosensitivity across the rat cervical spinal cord. Int J Radiat Oncol Biol Phys 2005;61(2):543–51, http://dx.doi.org/10.1016/j.ijrobp.2004.10.018.
[22] Korevaar EW, Habraken SJ, Scandurra D, Kierkels RG, Unipan M, Eenink MG, et al. Practical robustness evaluation in radiotherapy – a photon and proton-proof alternative to PTV-based plan evaluation. Radiother Oncol 2019;141:267–74, http://dx.doi.org/10.1016/j.radonc.2019.08.005.
[23] Wedenberg M, Toma-Dasu I. Disregarding RBE variation in treatment plan comparison may lead to bias in favor of proton plans. Med Phys 2014;41(9):091706, http://dx.doi.org/10.1118/1.4892930.
[24] Paganetti H. Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer. Phys Med Biol 2014;59(22):R419–72, http://dx.doi.org/10.1088/0031-9155/59/22/r419.
[25] Saager M, Peschke P, Brons S, Debus J, Karger CP. Determination of the proton RBE in the rat spinal cord: is there an increase towards the end of the spread-out Bragg peak? Radiother Oncol 2018;128(1):115–20, http://dx.doi.org/10.1016/j.radonc.2018.03.002.
[26] Cuaron JJ, Chang C, Lovelock M, Higginson DS, Mah D, Cahlon O, et al. Exponential increase in relative biological effectiveness along distal edge of a proton bragg peak as measured by deoxyribonucleic acid double-strand breaks. Int J Radiat Oncol * Biol * Phys 2016;95(1):62–9, http://dx.doi.org/10.1016/j.ijrobp.2016.02.018.
[27] Lühr A, von Neubeck C, Krause M, Troost EGC. Relative biological effectiveness in proton beam therapy – current knowledge and future challenges. Clin Transl Radiat Oncol 2018;9:35–41, http://dx.doi.org/10.1016/j.ctro.2018.01.006.
[28] Blanchard P, Wong AJ, Gunn GB, Garden AS, Mohamed AS, Rosenthal DI, et al. Toward a model-based patient selection strategy for proton therapy: external validation of photon-derived normal tissue complication probability models in a head and neck proton therapy cohort. Radiother Oncol 2016;121(3):381–6, http://dx.doi.org/10.1016/j.radonc.2016.08.022.
[29] Dutz A, Lühr A, Agolli L, Troost EGC, Krause M, Baumann M, et al. Development and validation of NTCP models for acute side-effects resulting from proton beam therapy of brain tumours. Radiother Oncol 2019;130:164–71, http://dx.doi.org/10.1016/j.radonc.2018.06.036.
[30] Pedersen J, Flampouri S, Bryant C, Liang X, Mendenhall N, Li Z, et al. Cross-modality applicability of rectal normal tissue complication probability models from photon- to proton-based radiotherapy. Radiother Oncol 2020;142:253–60, http://dx.doi.org/10.1016/j.radonc.2019.09.017.