Radiotherapy and Oncology 125 (2017) 520–525

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Proton therapy in oropharynx

The impact of treatment accuracy on proton therapy patient selection for oropharyngeal cancer patients

Tine Arts [a,*], Sebastiaan Breedveld [a], Martin A. de Jong [b], Eleftheria Astreinidou [b], Lisa Tans [a], Fatma Keskin-Cambay [a], Augustinus D.G. Krol [b], Steven van de Water [a], Rik G. Bijman [a], Mischa S. Hoogeman [a]

[a] Department of Radiation Oncology, Erasmus MC Cancer Institute, Rotterdam; and [b] Department of Radiation Oncology, LUMC, Leiden, The Netherlands

ARTICLE INFO

Article history:
Received 30 May 2017
Received in revised form 22 September 2017
Accepted 23 September 2017
Available online 23 October 2017

Keywords:
Proton therapy
Head and neck cancer
Oropharyngeal cancer
IMRT
IMPT
Robust optimization

ABSTRACT

Background and purpose: The impact of treatment accuracy on NTCP-based patient selection for proton therapy is currently unknown. This study investigates this impact for oropharyngeal cancer patients.

Materials and methods: Data of 78 patients was used to automatically generate treatment plans for a simultaneously integrated boost prescribing 70 Gy_{RBE}/54.25 Gy_{RBE} in 35 fractions. IMRT treatment plans were generated with three different margins; intensity modulated proton therapy (IMPT) plans for five different setup and range robustness settings. Four NTCP models were evaluated. Patients were selected for proton therapy if NTCP reduction was ≥10% or ≥5% for grade II or III complications, respectively.

Results: The degree of robustness had little impact on patient selection for tube feeding dependence, while the margin had. For other complications the impact of the robustness setting was noticeably higher. For high-precision IMRT (3 mm margin) and high-precision IMPT (3 mm setup/3% range error), most patients were selected for proton therapy based on problems swallowing solid food (51.3%) followed by tube feeding dependence (37.2%), decreased parotid flow (29.5%), and patient-rated xerostomia (7.7%).

Conclusions: Treatment accuracy has a significant impact on the number of patients selected for proton therapy. Therefore, it cannot be ignored in estimating the number of patients for proton therapy.

Radiation therapy (RT) combined with chemotherapy is frequently used to treat patients with head and neck cancer. RT is also associated with acute and late side effects that deteriorate quality of life (QoL) [1]. Intensity modulated proton therapy (IMPT) is a promising approach to reduce these adverse effects [2]. However, costs of IMPT exceed those of photon intensity modulated radiation therapy (IMRT) and world-wide IMPT capacity is limited. Therefore, IMPT should be applied to patients who are expected to benefit most.

Langendijk et al. proposed a model-based approach to select patients for proton therapy based on a reduction in normal tissue complication probability (ΔNTCP) calculated from a photon and a proton treatment plan. If ΔNTCP exceeds a pre-defined threshold level, e.g. 10% or 5% for a grade II or grade III complication respectively, IMPT is the treatment of choice [3]. This methodology gives rise to various concerns. One is that normal-tissue sparing also depends on the extra volume irradiated to mitigate errors in patient setup and proton range [4]. The impact of uncertainties and the measures to mitigate them, varies for IMRT and IMPT due to the physical differences between photons and protons. So while in photon therapy treatment uncertainties are typically compensated using safety margins, in IMPT they are increasingly dealt with using robust optimization. Recently, van der Voort et al. derived robustness recipes yielding the setup and range robustness settings for given distributions of systematic and random setup errors and systematic range errors [5]. However, the robustness settings that need to be used depend on the image-guidance procedures that will be applied. In addition, the IMRT margins are subject to change due to advances in image-guidance procedures. The aim of this study was to identify the impact of treatment accuracy on model-based IMPT patient selection for oropharyngeal cancer patients. To this purpose, IMRT and IMPT plans were automatically generated with various margins and robustness settings and the impact on patient selection was investigated for four IMRT-derived NTCP models for xerostomia, dysphagia, and tube feeding.

* Corresponding author at: Department of Radiation Oncology, Erasmus MC Cancer Institute, PO Box 5201, 3008 AE Rotterdam, The Netherlands.
E-mail address: t.arts-3@umcutrecht.nl (T. Arts).

https://doi.org/10.1016/j.radonc.2017.09.028
0167-8140/

Materials and methods

Patient group and treatment plan generation

Anonymized CT data and structure sets of 78 consecutive oropharyngeal patients were used, of whom 24 patients were previously treated at the Leiden University Medical Center (LUMC) and 54 patients at the Erasmus MC Cancer Institute. Characteristics are listed in Table 1. All patients were planned using a simultaneously integrated boost scheme prescribing 70 Gy_{RBE} to the primary tumour and pathological lymph nodes (LUMC) or levels with pathological lymph nodes (Erasmus MC) and 54.25 Gy_{RBE} to the elective nodal areas in 35 fractions. For IMPT, "minimax" robust optimization (see section "Margins and robustness settings") was applied to the unmodified clinical target volumes (CTVs). For IMRT CTVs were expanded to planning target volumes (PTVs). The planning goal was that ≥95% of the prescribed dose should be received by ≥98% of the PTV (IMRT) or CTV of the worst-case robustness scenario (IMPT). A constant radiobiological effectiveness (RBE) of 1.0 and 1.1 was assumed for the IMRT and IMPT plans respectively [6]. All plans were generated using Erasmus-iCycle, an in-house developed optimizer [7,8]. This optimizer allows to efficiently generate treatment plans for a large cohort of patients in a fully automated fashion. Input for this optimizer is a user-defined wish-list, composed of constraints and prioritized objectives, where each objective is assigned a certain goal. Based on this wish-list, the multi-criterial optimizer optimizes the objectives one-by-one according to the set priorities. In contrast to the objectives, the constraints have to be met at all times. Separate wish-lists, but with similar intent, were used for IMRT and IMPT plans (see Supplementary material) [9,10]. Both wish-lists were constructed based on the same treatment objectives. However, due to the different physical characteristics between photons and protons, the used wish-lists are not identical. The wish-lists were designed in close collaboration with radiation oncologists. For the IMRT plans, we used a 23 equi-angular beam arrangement to simulate volumetric arc therapy (VMAT) dose distributions [9]. The dose was computed in CT-resolution ($0.98 \times 0.98 \times 2.5$ mm^3). For IMPT we used three equi-angular beams at 60°, 180° and 300°, as suggested by literature [11]. Available proton energies ranged from 69 to 250 MeV with corresponding spot widths ranging from 3.8 to 6.0 mm sigma (in air at the isocentre), respectively. To irradiate superficially located target regions, we assumed that a range shifter of 57 mm water equivalent thickness could be inserted during the delivery of a field. Pencil beams were selected and optimized using the resampling method described by van de Water et al. [12]. Final dose calculation was performed on a $2 \times 2 \times 2$ mm^3 grid and interpolated to CT-resolution. In case of minor violations in target coverage (<1%) after final dose calculation, the dose distribution was rescaled to again fulfil the constraint $V_{95\%} \geq 98\%$.

CT artefacts were present in 45 patients due to metal dental artefacts (e.g. fillings). The artefacts may impact IMPT treatment plan generation and subsequently the NTCP values. Therefore, IMPT treatment plans were generated before and after artefact reduction (Metal Deletion Technique v1.1, Revision Radiology) for five patients with the most severe artefacts.

Table 1
Patient and tumour characteristics.

Characteristics		Number	%
Sex	Male	58	74
	Female	20	26
Age	<65	47	60
	>65	31	40
T-classification	T1	4	5
	T2	42	54
	T3	12	15
	T4	20	26
Bilateral neck irradiation	Yes	71	91
	No	7	9
Weight loss	None	59	75
	Moderate	17	22
	Severe	2	3
Accelerated radiotherapy	Yes	38	49
	No	40	51
Radiotherapy plus Cetuximab	Yes	14	18
	No	64	82
Chemoradiation	Yes	22	28
	No	56	72

Margins and robustness settings

For IMRT, the CTV was isotropically expanded with a 0, 3, or 5 mm margin to account for geometrical uncertainties with a 5 mm retraction under the patient's skin [13]. For IMPT, robust optimization was used to account for uncertainties using setup robustness and range robustness. Nine scenarios were included: setup errors in the positive and negative direction along three axes (six scenarios), positive and negative range errors (two scenarios) and one nominal scenario (no errors). Erasmus-iCycle includes these nine scenarios simultaneously using a "minimax" optimization [14,15], and optimizes the worst-case scenario for each objective. Fractionation is not considered directly, but similar to margins robustness recipes can be used to determine for fractionated treatments the settings needed to ensure adequate CTV coverage in patients for given random and systematic error distributions [5]. Setup error scenarios were simulated by laterally shifting the pencil beams. The range error scenarios were generated by altering the proton energy. Hereto, we transformed the range error into an equivalent energy adjustment for each spot. The IMRT margins and IMPT robustness settings are summarized in Table 2. IMRT plans with 0 mm margins and IMPT plans with 0 mm setup robustness (SR = 0 mm) and 0% relative range robustness (RR = 0%) were included for a baseline comparison between IMRT and IMPT.

Plan evaluation

All IMRT plans were evaluated for meeting the clinical target goals ($V_{95\%} \geq 98\%$) for the low-dose as well as the high dose PTV and $V_{107\%} \approx 2\%$ and $V_{110\%} \approx 0\%$ for the high dose PTV. For IMPT, we evaluated the same parameters but then for the CTVs of the nominal and error scenarios. The dose to organs at risk (OARs) was checked for outliers and all IMRT and nominal IMPT dose distributions were evaluated visually.

NTCP models

Published NTCP models recently discussed for IMPT patient selection in the Netherlands were used to compare IMRT and IMPT plans and assess the impact of margins and robustness settings on

Table 2
Used margins and robustness settings for IMRT and IMPT plans. IMPT robustness settings were first sorted to setup robustness and second to range robustness as setup robustness has a larger impact on OAR dose than range robustness [9].

IMRT	IMPT	
Margin (mm)	Setup Robustness (SR) (mm)	Range Robustness (RR) (%)
0	0	0
3	3	3
5	3	5
	5	3
	5	5

Abbreviations: IMRT = intensity modulated radiation therapy; IMPT = intensity modulated proton therapy; OAR = organ at risk.

patient selection. Table 3 lists the four models and their properties [16–19]. Similarly as in Jakobi et al. [20] the organ receiving the lowest dose of two paired organs, such as the parotid gland, was appointed as the contralateral organ. For the model for patient-rated xerostomia, the baseline xerostomia score (0 = none or 1 = a bit) was not recorded for our patient group. Therefore we randomly assigned 30% of our patient group with baseline xerostomia = 1 and 70% with baseline xerostomia = 0, which is the ratio found in the patient training set in the article of Beetz et al. [16]. For the model for decreased parotid flow of Dijkema et al. [17] the left and right parotid were handled separately when calculating and comparing NTCP values, meaning we compared the left parotid glands mutually ($\Delta NTCP_{left} = NTCP_{left,IMRT} - NTCP_{left,IMPT}$) and the right parotid glands mutually ($\Delta NTCP_{right} = NTCP_{right,IMRT} - NTCP_{right,IMPT}$). The final ΔNTCP was then the maximum of these, i.e. $\max(\Delta NTCP_{left}, \Delta NTCP_{right})$.

For plan comparison the NTCP values of the IMPT plans were subtracted from those of the IMRT plans, resulting in a ΔNTCP. Decision making in favour of IMPT was made when the ΔNTCP threshold (10% or 5% for a grade II or grade III complication, respectively) was exceeded. These thresholds had recently been set by the Dutch Society for Radiation Therapy and Oncology. In addition, we investigated the impact of the ΔNTCP threshold on the number of patients selected for IMPT.

Contouring organs at risk

The OARs considered in the NTCP models were the parotid glands, supraglottic larynx, superior and inferior constrictor muscle (MCS and MCI respectively), and the cricopharyngeal muscle (MCP). Delineation followed published guidelines [16,21–24]. The supraglottic larynx was delineated for 17 patients first. Atlas-based auto-segmentation (Elekta AB, CMS software, version 0.63, St. Louis) was used to propagate the supraglottic larynx delineations to the remaining patients.

Overlap in selection

The NTCP models proposed for IMPT patient selection partially overlap in type of complication and input parameters (e.g. the mean dose of OARs). The latter may lead to an overlap in selected patients between two NTCP models. To determine this overlap we calculated:

$$Overlap\ model\ 1\ with\ model\ 2 = \frac{\#SelectedPtsModel1 \cap \#SelectedPtsModel2}{\#SelectedPtsModel1} \times 100\%,$$

where #*SelectedPtsModel*1 and #*SelectedPtsModel*2 are the number of patients selected for IMPT by models 1 and 2, respectively. It gives the percentage of patients selected by model 2 from the group of patients already selected by model 1. The overlap was calculated for a 3 mm margin and a robustness setting of SR = 3 mm/RR = 3%, assuming high-precision IMRT and IMPT [5,25].

Results

Plan quality

For all IMRT plans, $V_{95\%}$ was above 98% for the high-dose and low-dose PTV and $V_{107\%}$ was below 2% and $V_{110\%}$ = 0% for the high-dose PTV. For the IMPT plans, $V_{95\%}$ was above 98% for the high-dose as well as the low-dose CTV error scenarios. The SR = 0 mm/RR = 0% treatment plans of 8 patients had to be rescaled as the $V_{95\%}$ was slightly lower than 98% (range 97.3–97.7%) after recalculation on the fine dose grid. For 7 treatment plans (4 patients), the $V_{107\%}$ in the worst-case error scenario was above 2% (range: 8.0–10.6%) after recalculation on the fine dose grid. As in the nominal scenario the $V_{107\%}$ was below 2% and the $V_{110\%}$ was 0% in all scenarios, those plans were rendered clinically acceptable.

For the five patients with the most severe CT artefacts the largest difference in NTCP before and after artefact reduction was seen for patient-rated xerostomia (average: 0.3%; range: −0.4 to 1.2%). Given these small differences, artefact reduction was not performed for any of the patients.

Patients selected for IMPT per model

The percentage of patients selected for IMPT as a function of margin and robustness setting is shown in Fig. 1 for each NTCP model. In general, the percentage of patients selected for IMPT decreases with increasing robustness setting for a given margin. Similarly, Fig. 1 shows that for a given robustness setting the percentage of patients selected for IMPT decreases with decreasing margin. Concerning tube feeding dependence and problems swallowing solid food, patients are selected for IMPT even in case of the hypothetical use of a 0 mm margin compared to nonzero robustness. The degree of robustness, however, has little impact on patient selection for tube feeding dependence. For problems swallowing solid food the impact of the degree of robustness setting is markedly higher as well as for patient-rated xerostomia for a margin of 5 mm. For a 3 mm margin and robustness setting of SR = 3 mm/RR = 3%, most patients are selected for IMPT based on problems swallowing solid food (51.3%), followed by tube feeding dependence (37.2%) and decreased parotid flow (30.8%).

Table 3
NTCP models used for plan comparison.

NTCP model	Grade	Endpoint	Parameters
Wopken et al. [19]	III	Tube feeding dependence after six months	Mean dose of the superior PCM, inferior PCM, contralateral parotid and cricopharyngeal muscle Advanced T-stage Weight loss (moderate/severe) Accelerated radiotherapy Chemoradiation Radiotherapy plus cetuximab
Dijkema et al. [17]	II	<25% Parotid flow for individual parotid gland after 1 year	Mean dose in parotid glands
Christianen et al. [18]	II	Problems swallowing solid food assessed with the EORTC QLQ-H&N35 questionnaire	Mean dose superior PCM and supraglottic larynx Age
Beetz et al. [16]	II	Moderate-to-severe patient-rated xerostomia after six months assessed by the EORTC QLQ-H&N35 questionnaire	Mean dose contralateral parotid gland Baseline xerostomia score

Abbreviations: NTCP = normal tissue complication probability; PCM = pharyngeal constrictor muscle.

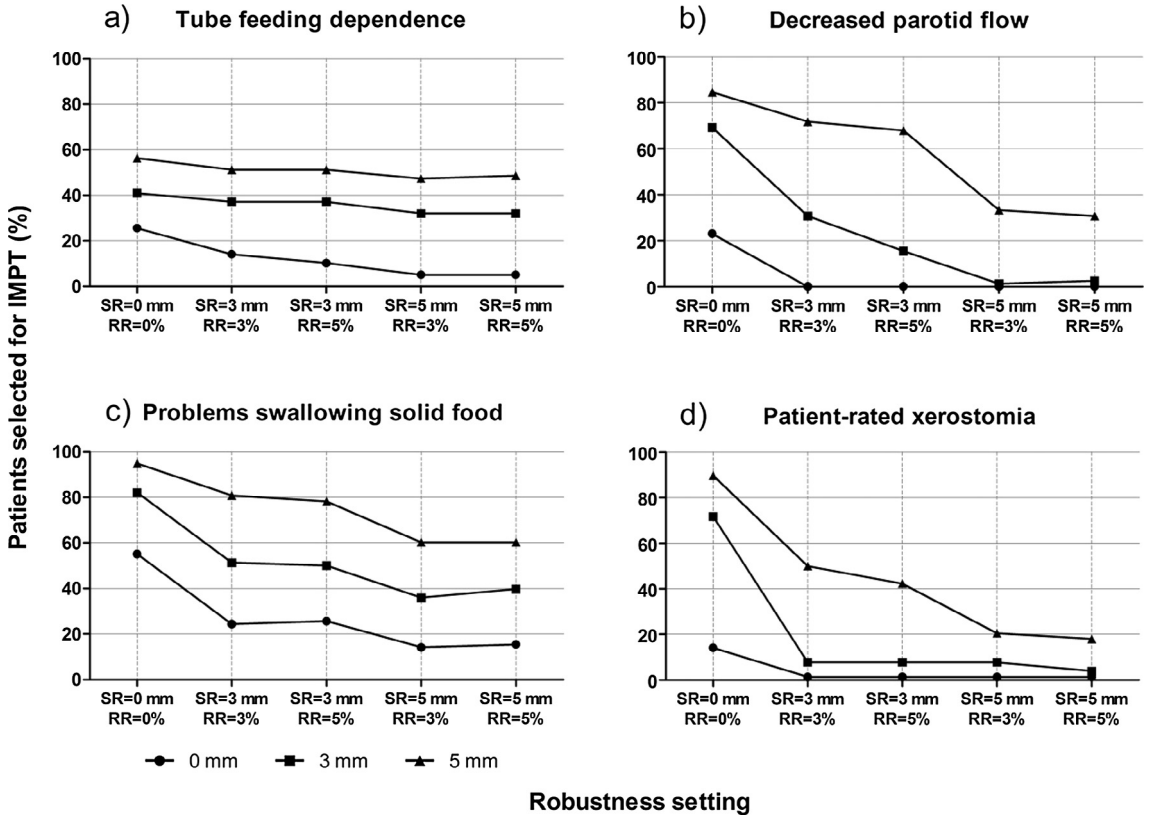

Fig. 1. (a–d) Percentage of patients selected for IMPT (intensity modulated proton therapy) as a function of margins and robustness settings for each of the four models included in the comparison. A ΔNTCP (normal tissue complication probability) threshold of 10% for the grade II models and 5% for the grade III model is used. SR = setup robustness, RR = range robustness.

Patient-rated xerostomia contributed only with 7.7% (see Supplementary materials for absolute NTCP values and their standard deviations). For the models for tube feeding dependence, decreased parotid flow, and patient-rated xerostomia, it is noteworthy to mention that the number of selected patients is relatively stable if IMRT and IMPT plans with similar accuracy, i.e. with similar margin and setup robustness, are compared.

Overlap in selection

Table 4 shows the overlap between the NTCP models given a 3 mm margin and a robustness setting of SR = 3 mm/RR = 3%. For tube feeding dependence 29 patients were selected. Of those patients 55.2% was also selected for decreased parotid flow. Considering the three models with the highest number of patients selected, the largest overlap is 66.7% indicating that all three models contribute for a great portion independently to the number of patients selected for proton therapy. Also, for the six patients selected based on patient-rated xerostomia, there is no other model that selects all these six patients. For high precision IMRT (3 mm margin) as well as IMPT (SR = 3 mm/RR = 3%) the union of the percentage of patients selected for IMPT by all four NTCP models is 77%.

Impact of ΔNTCP threshold

Fig. 2 shows the percentage of patients selected for IMPT as a function of ΔNTCP threshold values. The models for patient-rated xerostomia, problems swallowing solid food and decreased parotid flow show the same trend. The model for tube feeding dependence shows the sharpest decrease in patient selection for IMPT at the lowest ΔNTCP threshold values. For higher ΔNTCP thresholds the majority of patients will be selected for problems swallowing solid food.

Discussion

In this study IMRT as well as IMPT plans were automatically generated with various margins and robustness settings for 78 patients in order to investigate the impact of treatment accuracy on patient selection for proton therapy. Based on the results we conclude that treatment accuracy cannot be ignored in estimating the number of patients that will be selected for proton therapy. Improvements in the accuracy of IMPT, IMRT, or both, for example by implementing improved image guidance techniques can change patient selection for IMPT. However, if we assume that the treatment-related accuracy of IMRT and IMPT is equivalent and

Table 4
Overlap between the NTCP models given in percentages. Numbers indicate the percentage overlap of the model in the first column with the model in the first row.

	Tube feeding dependence	Decreased parotid flow	Problems swallowing solid food	Patient-rated xerostomia
Tube feeding dependence (29)		55.2%	58.6%	3.4%
Decreased parotid flow (24)	66.7%		58.3%	4.4%
Problems swallowing solid food (40)	42.5%	35.0%		5.0%
Patient-rated xerostomia (6)	16.7%	16.7%	33.3%	

Abbreviations: NTCP = normal tissue complication probability.

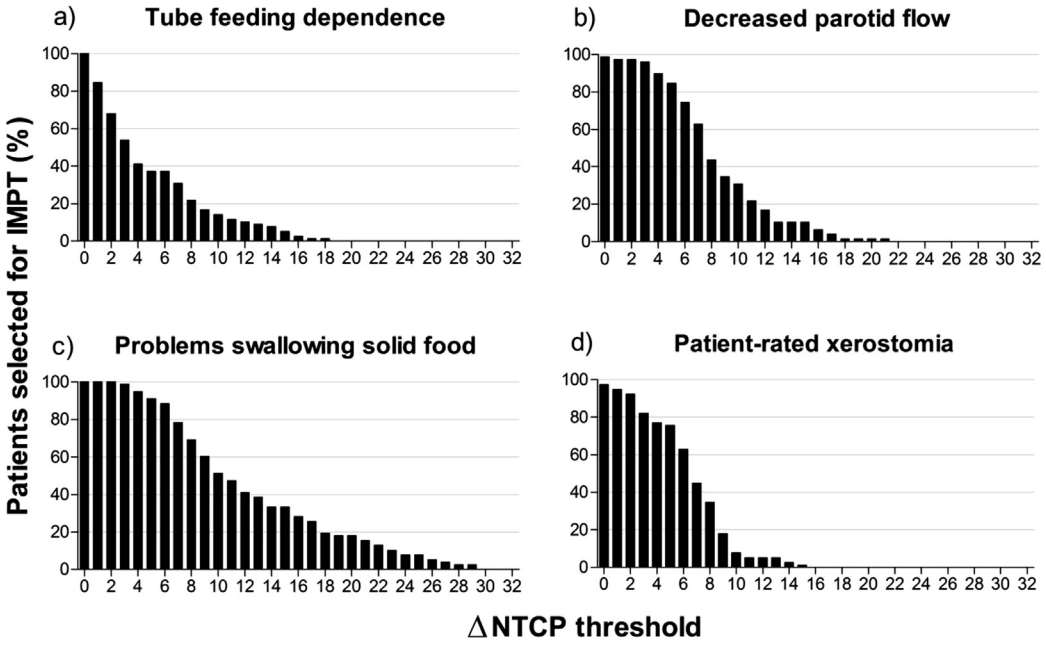

Fig. 2. (a–d) Percentage of patients selected for IMPT (intensity modulated proton therapy) as a function of the ▫ NTCP (normal tissue complication probability) threshold for all four NTCP models. Results are based on IMRT (intensity modulated radiation therapy) plans with a 3 mm margin and IMPT plans with SR = 3 mm/RR = 3%. All 78 patients were included. SR = setup robustness, RR = range robustness.

that similar margins and setup robustness settings result both in adequate CVT coverage, the number of patients selected is relatively invariant under treatment accuracy. The actual impact of treatment accuracy on patient selection depends on the parameters included in the models, but none of the four models remained unaffected. In our study, the degree of setup robustness had a larger impact on patient selection than range robustness. This is in agreement with van de Water et al. [10] who showed that the degree of setup robustness had the biggest impact on NTCP.

Our results showed that the impact of metal dental artefacts on NTCP is very small and could therefore be ignored in treatment planning for patient selection. We believe that this can be explained by the fact that the artefacts are present in only a few CT slices, while the organs at risk usually extend beyond those few slices or do not overlap with those slices. We would like to state that in the deliverable treatment plan the metal dental artefacts have to be accounted for.

The generation of a treatment plan usually implies a trade-off between the coverage of the target volume and sparing of OARs. For treatment plans that are Pareto optimal [26], dose improvements for a certain OAR will automatically lead to a worsening of the dose to the target volume or other OARs. In our study, automated planning was employed using prioritized optimization based on pre-defined wish-lists to generate Pareto optimal treatment plans. This guarantees the same trade-off between planning objectives across all patients contributing to consistent results on the impact of treatment accuracy on proton therapy patient selection. To what extent wish-lists with alternative prioritizations in OAR sparing and target coverage or patient-specific prioritizations may impact patient selection would be interesting to investigate.

Obviously there are inherent differences between treatment planning for IMRT and IMPT, as IMRT uses PTV margins while robust optimization was used for IMPT. To that end, it is not straightforward to compare a specific margin to a specific robustness setting for a fractionated treatment with similar accuracy. van der Voort et al. derived for the first time robustness recipes for IMPT [5]. Their results can be used to find for a given treatment accuracy in a fractioned treatment, the robustness settings that would lead to an adequate treatment in a population of head and neck cancer patients. Still, it would be an interesting topic of future research to compare IMRT and IMPT treatment plans that are both generated using robust optimization.

In clinical practice, the Erasmus-iCycle 23-beam IMRT treatment plans are reconstructed automatically in our clinical treatment planning system (Monaco Elekta AB, Stockholm) to deliverable VMAT plans [9]. The same approach is foreseen for the IMPT treatment plan. To avoid inducing additional biases resulting from the reconstruction, we limited patient selection to the treatment plans generated in Erasmus-iCycle. The final decision to treat a patient with proton therapy should be based on a clinically deliverable plan. Therefore, minor differences in patient selection are expected between the automated approach used in this study and the final selection based on clinically deliverable plans.

A limitation of our study was the absence of baseline xerostomia data, which is one of the parameters of the patient-rated xerostomia model. We randomly assigned baseline xerostomia to 30% of the patients included in this study. To analyse the impact of this approach, we also analysed the results assuming that all patients had baseline xerostomia or had no baseline xerostomia. We found that the baseline score did hardly impact our findings.

Another limitation is that the NTCP models used in this study were derived from photon treatments, which may result in reduced accuracy in predicting complications for IMPT. In a recent study of Blanchard et al., however, it was demonstrated that photon-based NTCP models were still valid in a cohort of proton-treated head and neck cancer patients [27]. The impact of the accuracy of the NTCP models themselves on the accuracy of IMPT patient selection is topic of current research.

Fig. 1 shows that for problems swallowing solid food, decreased parotid flow, and tube feeding dependence the number of patients selected for IMPT sometimes slightly increased for increasing robustness. This is most likely a result of the prioritized optimization method. An increase in robustness setting may limit the sparing of a highly prioritized OAR. This can give the optimizer more freedom to decrease the dose in lower prioritized OARs. Depending

on which OARs benefit and which not, it can result in a decrease in NTCP for increased robustness setting.

Conclusion

This study shows that treatment accuracy cannot be ignored in estimating the number of patients selected for proton therapy based on comparative treatment planning and NTCP evaluations. It also shows that IMRT and IMPT image-guidance techniques should be up-to-date, otherwise the patient selection is based on treatment accuracy and not on the physical properties of the radiation applied.

Conflict of interest

Erasmus MC Cancer institute has research collaborations with Elekta AB, Stockholm, Sweden, Accuracy Inc, Sunnyvale, USA, and Varian Medical Systems, Palo Alto, USA.

Role of the funding source

The funding source had no involvement in this study.

Acknowledgements

This research was in part funded by Holland Proton Therapy Center and by Erasmus MC Mrace-efficiency grant 2014-14209.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.radonc.2017.09.028.

References

[1] Bhide SA, Newbold KL, Harrington KJ, et al. Clinical evaluation of intensity-modulated radiotherapy for head and neck cancers. Br J Radiol 2012;85:487–94.
[2] Van de Water TA, Lomax AJ, Bijl HP, et al. Potential benefits of scanned intensity-modulated proton therapy versus advances photon therapy with regard to sparing of the salivary glands in oropharyngeal cancer. Int J Radiat Oncol Biol Phys 2011;79:1216–24.
[3] Langendijk JA, Lambin P, De Ruysscher D, et al. Selection of patients for radiotherapy with proton aiming at reduction of side-effects: the model-based approach. Radioth Oncol 2013;107:267–73.
[4] Liu W, Frank SJ, Li X, et al. Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers. Med Phys 2013;40:051711.
[5] Van der Voort S, van de Water S, Perkó Z, et al. Robustness recipes for minimax robust optimization in intensity modulated proton therapy for oropharyngeal cancer patients. Int J Radiat Oncol Biol Phys 2016;95:163–70.
[6] Paganetti I, Niemierko A, Ancukiewicz M, et al. Relative biological effectiveness (RBE) values for proton beam therapy. Int J Radiat Oncol Biol Phys 2002;53:407–21.
[7] Breedveld S, Storchi PR, Voet PW, et al. iCycle: integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. Med Phys 2012;39:951–63.
[8] Breedveld S, Storchi PR, Heijmen BJ. The equivalence of multi-criteria methods for radiotherapy plan optimization. Phys Med Biol 2009;54:7199–209.
[9] Voet PW, Dirkx ML, Breedveld S, et al. Towards fully automated multicriterial plan generation: a prospective clinical study. Int J Radiat Oncol Biol Phys 2013;85:866–72.
[10] Van de Water S, van Dam I, Schaart DR, et al. The price of robustness; impact of worst-case optimization on organ-at-risk in intensity-modulated proton therapy for oropharyngeal cancer patients. Radiother Oncol 2016;120:56–62.
[11] Van de Water TA, Lomax AJ, Bijl HP, et al. Using a reduced spot size for intensity modulated proton therapy potentially improves salivary gland-sparing in oropharyngeal cancer. Int J Radiat Oncol Biol Phys 2012;82:e313–9.
[12] Van de Water S, Kraan AC, Breedveld S, et al. Improved efficiency of multi-criterial IMPT treatment planning using iterative resampling of randomly placed pencil beams. Phys Med Biol 2013;58:6969–83.
[13] Astreinidou E, Bel A, Raaijmakers CPJ, et al. Adequate margins for random setup uncertainties in head-and-neck IMRT. Int J Radiat Oncol Biol Phys 2005;61:938–44.
[14] Chen W, Unkelbach J, Trofimov A, et al. Including robustness in multi-criteria optimization for intensity modulated proton therapy. Phys Med Biol 2012;57:591–608.
[15] Fredriksson A, Forsgren A, Hårdemark B. Minimax optimization for handling range and setup uncertainties in proton therapy. Med Phys 2011;38:1672–84.
[16] Beetz I, Schilstra C, van der Schaaf A, 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;105:101–6.
[17] Dijkema T, Raaijmakers CP, Ten Haken RK, et al. Parotid gland function after radiotherapy: the combined Michigan and Utrecht experience. Int J Radiat Oncol Biol Phys 2010;78:449–53.
[18] Christianen ME, Schilstra C, Beetz I, et al. Predictive modelling for swallowing dysfunction after primary (chemo)radiation: results of a prospective observational study. Radiother Oncol 2012;105:107–14.
[19] Wopken K, Bijl HP, van der Schaaf A, 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. Radioth Oncol 2014;113:95–101.
[20] Jakobi A, Bandurska-Lugua A, Stützer K, et al. Identification of patient benefit from proton therapy for advanced head and neck cancer patients based on individual and subgroup normal tissue complication probability analysis. Int J Radiat Oncol Biol Phys 2015;92:1165–74.
[21] Beale T, Madani G. Anatomy of the salivary glands. Semin Ultrasound CT MR 2006;27:436–9.
[22] Brouwer CL, Steenbakkers RJ, Bourhis J, et al. CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC, NTG, NCRI, NTG oncology and TROG consensus guides. Radiother Oncol 2015;117:83–90.
[23] Haderlein M, Semrau S, Otto O, et al. Dose-dependent deterioration of swallowing function after induction chemotherapy and definitive chemoradiotherapy for laryngopharyngeal cancer. Strahlenther Onkol 2014;190:192–8.
[24] Heukelom J, Hamming O, Bartelink H, et al. Adaptive and innovative radiation treatment for improving cancer treatment outcome (ARTFORCE); a randomized controlled phase II trial for individualized treatment of head and neck cancer. BMC Cancer 2013;13:84.
[25] Chen AM, Yu Y, Daly ME, et al. Long-term experience with reduced planning target volume margins and intensity-modulated radiotherapy with daily image guidance for head and neck cancer. Head Neck 2014;36:1766–72.
[26] Miettinen K. Nonlinear multiobjective optimization. Boston: Kluwer Academic Publishers; 1999.
[27] Blanchard P, Wong AJ, Gunn GB, 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. Radioth Oncol 2016;121:381–6. https://doi.org/10.1016/j.radonc.2016.08.022.