First experience with model-based selection of head and neck cancer patients for proton therapy

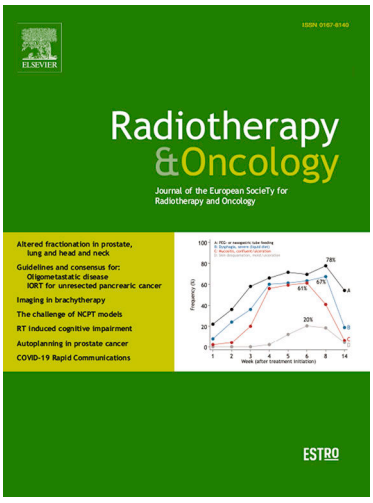
Tambas, Makbule; J H M Steenbakkers, Roel; van der Laan, Hans P; Wolters, Atje M; Kierkels, Roel G J; Scandurra, Dan; Korevaar, Erik W; Oldehinkel, Edwin; van Zon-Meijer, Tineke W H; Both, Stefan

Published in:
Radiotherapy and Oncology

DOI:
10.1016/j.radonc.2020.07.056

Publication date:
2020

Journal Pre-proofs



Original Article

First experience with model-based selection of head and neck cancer patients for proton therapy

Makbule Tambas, Roel J.H.M. Steenbakkers, Hans P. van der Laan, Atje M. Wolters, Roel G.J. Kierkels, Dan Scandurra, Erik W. Korevaar, Edwin Oldehinkel, Tineke W.H. van Zon-Meijer, Stefan Both, Johanna G.M. van den Hoek, Johannes A. Langendijk

PII: S0167-8140(20)30715-5
DOI: https://doi.org/10.1016/j.radonc.2020.07.056
Reference: RADION 8471

To appear in: *Radiotherapy and Oncology*

Received Date: 29 May 2020
Revised Date: 30 July 2020
Accepted Date: 31 July 2020

Please cite this article as: Tambas, M., J.H.M. Steenbakkers, R., van der Laan, H.P., Wolters, A.M., Kierkels, R.G.J., Scandurra, D., Korevaar, E.W., Oldehinkel, E., van Zon-Meijer, T.W.H., Both, S., van den Hoek, J.G.M., Langendijk, J.A., First experience with model-based selection of head and neck cancer patients for proton therapy, *Radiotherapy and Oncology* (2020), doi: https://doi.org/10.1016/j.radonc.2020.07.056

Title Page

First experience with model-based selection of head and neck cancer patients for proton therapy

Short running title: Model-based selection for proton in HNC

Makbule Tambas[a,*], Roel J.H.M. Steenbakkers[a], Hans P. van der Laan[a], Atje M. Wolters[a], Roel G.J. Kierkels[a,b], Dan Scandurra[a], Erik W. Korevaar[a], Edwin Oldehinkel[a], Tineke W.H. van Zon-Meijer[a], Stefan Both[a], Johanna G.M. van den Hoek[a], Johannes A. Langendijk[a]

[a]University of Groningen, University Medical Center Groningen, Department of Radiation Oncology, Hanzeplein 1, 9713 GZ Groningen, the Netherlands

[b]Radiotherapiegroep, Department of Radiation Oncology, Wagnerlaan 47, 6815 AD Arnhem / Deventer, the Netherlands.

***Corresponding Author**: Makbule Tambas

University of Groningen, University Medical Center Groningen, Department of Radiation Oncology, Hanzeplein 1, 9713 GZ Groningen, the Netherlands

e-mail: [¥]m.tambas@umcg.nl, telephone: +31 050 361 6161

Conflict of Interest: None

Funding: This work described in this paper has been granted by the European Union (Project number: 730983: INSPIRE-project).

Abstract

Purpose: In the Netherlands, head and neck cancer (HNC) patients qualify for intensity modulated proton therapy (IMPT) based on model-based selection (MBS). The aim of this study was to evaluate the first experience in MBS of HNC patients.

Methods: Patients who were subjected to MBS (Jan 2018 - Sep 2019) were evaluated. A VMAT plan was created for all patients with optimal sparing of organ at risks (OARs) in normal tissue complication probability (NTCP) models for a number of toxicities. An IMPT plan was created only for

those with NTCP difference (ΔNTCP) between VMAT and best-case scenario for proton (assuming 0 Gy dose for all OARs in IMPT plan) that exceeded any ΔNTCP-thresholds defined in Dutch National Indication Protocol. These patients qualified for a robust IMPT-plan creation with similar target doses and subsequent plan comparison.

Results: Of 227 patients, 141 (62%) qualified for plan comparison, of which 80 (35%) were eventually selected for proton therapy. Most patients were selected based on the ΔNTCP for dysphagia-related toxicities. The selection rate was higher among patients with advanced disease, pharyngeal tumors, and/or baseline complaints. A significant reduction in all OAR doses and NTCP values was obtained with IMPT compared with VMAT in both selected and non-selected patients, but more pronounced in patients selected for protons.

Conclusion: Model-based selection of patients with HNC for proton therapy is clinically feasible. Approximately one third of HNC patients qualify for protons and these patients have the highest probability to benefit from protons in terms of toxicity prevention.

Keywords: Proton therapy, head and neck cancer, patient selection, model-based selection.

Abbreviations

HNC, head and neck cancer; IMPT, intensity modulated proton therapy; MBS, based on model-based selection; OARs, organ at risks; NTCP, normal tissue complication probability; ΔNTCP, NTCP difference; NIPP, National Indication Protocol Proton therapy; TFD, tube feeding dependence; ΣΔNTCP, the summed risk reduction; PCM, pharyngeal constrictor muscle; PBS, pencil beam scanning.

Introduction

Compared to radiotherapy using photons, the superior beam properties of protons can be translated into two main strategies to obtain a clinical benefit: 1) dose escalation to improve local tumor control without increasing the risk of unacceptable radiation-induced toxicity, or 2) the sparing of healthy tissues with equivalent target dose to reduce the risk of radiation-induced side effects with similar tumor control. For the latter group of indications, the model-based approach was developed [1,2]. Model-based selection identifies patients who are expected to benefit most from protons. In the Netherlands, around 2,000 HNC patients are treated with radiotherapy each year and the expected proportion of patients qualifying for proton therapy is around 30-40% [2,3].

The model-based approach has been approved and accepted as an evidence-based method by the Dutch Health Care Institute (Zorginstituut Nederland) to select patients for proton therapy, when the aim is to reduce radiation-induced toxicity [4]. Patients selected according to this method receive full reimbursement for proton therapy.

In November 2017, the first National Indication Protocol Proton therapy (NIPP) for model-based selection of head and neck cancer patients (HNC) was accepted and implemented in January 2018 at our hospital [2]. The aim of this study was to describe the first experience with model-based selection of HNC patients for proton therapy according to the first version of the NIPP.

Materials and Methods

Patients

This is a prospective cohort study, including HNC patients who were subjected to the model-based selection procedure from January 2018 to September 2019, based on the first version of the NIPP. Only tumors originating from the oral cavity, pharynx, and larynx were considered for model-based selection. All patients were treated with definitive radiotherapy, with or without systemic treatment

(i.e., weekly cisplatin or cetuximab), with a prescribed target dose of 70 Gy and 54.25 Gy in 35 fractions to the high risk and prophylactic lymph nodal regions, respectively, using a simultaneous integrated boost technique. For photon and proton plans, the same dose and fractionation schedules were prescribed [5].

Delineation of the Organs-at-Risk (OARs)

The OARs delineation was performed according to the international consensus guidelines for CT-based delineation of OARs in the head and neck region [6]. The tumor border and its invasion into the OARs were delineated on the simulation CT using MRI and PETC/CT imaging of the patients and checked by a multidisciplinary team including radiation oncologists, radiologists and nuclear medicine specialists. Contouring of the OARs were not adjusted in case of overlap with target volumes, meaning that the D_{mean} as mentioned in the paper reflects the actual dose in the entire OAR, including the part that overlapped with the target. In some cases, dose reduction in these OARs using protons was limited as underdosage in the target volume was not permitted.

National Indication Protocol for Proton therapy (NIPP)

The content and background of the model-based selection of patients for proton therapy has previously been described by Langendijk et al. [1,2] and the workflow is shown in Figure 1. For HNC patients, the NTCP models and ΔNTCP threshold values to be used patient selection for proton therapy were based on a detailed literature review and expert consensus in the Dutch National Platform for proton therapy. Three NTCP-models were selected: 1) patient-rated moderate to severe xerostomia, derived from Beetz et al. [7]; 2) physician-rated grade ≥ II dysphagia derived from Christianen et al. [8]; and 3) for tube feeding dependence (TFD) derived from Wopken et al. [9] six months after radiotherapy. These three models were subjected to external validation and updated in an independent patient cohort (Table 1).

According to the NIPP, patients qualify for proton therapy in case the ΔNTCP ≥ 10% for grade ≥ II and/or ≥ 5%, for grade ≥ III side effects, respectively. Patients also qualify for protons if the summed risk reduction (ΣΔNTCP) for grade ≥ II side effects is 15% or higher. The predictors of the NTCP-models and related ΔNCTP thresholds are shown in Table 1.

Model-based selection procedure

The model-based selection procedure consists of 4 steps:

Step 1: Model-based optimized VMAT-plan

In each patient, a VMAT plan was made in which at least 98% of the Planning Target Volume (PTV) was covered by 95% of the prescribed dose, while dose to the 2% of PTV was kept below 107%. We used model-based optimization aimed at reducing the dose in the OARs that were included as predictors in the NTCP-models.

After adequate coverage of PTVs, OARs were prioritized for optimization in the following order: 1) both parotid glands; 2) swallowing structures (pharyngeal constrictor muscle (PCM) superior, PCM inferior, cricopharyngeal muscle); and 3) oral cavity [5,6]. When there was an overlap between target volumes and these OARS, dose optimization was guided based on the NTCP-model gradients, i.e., the NTCP gain (%) for each OAR if their D_{mean} was reduced by 1 Gy. Based on the NTCP-gain % per Gy, dose optimization was performed by prioritizing the OARs with the highest NTCP gain per Gy. In addition, the dose to the other OARs was decreased as much as possible without compromising the dose to the prioritized structures.

The RayStation treatment planning system (v6.1 and v8, RaySearch Laboratories AB, Stockholm, Sweden) was used for all VMAT plans. The VMAT plans consisted of 6 MV energy, dual arcs, and dose grid resolution of 3 mm^3. The final photon dose was calculated using a collapsed cone dose engine.

Step 2: Preselection Tool

To select patients who qualified for the creation of an IMPT plan and subsequent plan comparison, we created a calculation algorithm which we used as a preselection tool. The DVH and non-DVH data of the patient relevant for the NTCP models were entered into this preselection tool after completing the VMAT plan. The preselection tool produced two NTCP-profiles, i.e., the calculated complication risks for xerostomia, dysphagia and TFD: 1) for the clinically prepared VMAT plan and 2) for the best-case scenario IMPT plan (assuming a dose of zero Gy for all OARs). If the ΔNTCP-value between these two profiles exceeded at least one of the four ΔNTCP-thresholds (Table 1), then patients qualified for the creation of a model-based IMPT plan and subsequent plan comparison.

Step 3: Model-based optimized IMPT-plan

For each patient who was selected for an IMPT plan creation and subsequent plan comparison based on the PST, a pencil beam scanning (PBS) intensity-modulated proton plan (IMPT) was created with robust optimization in RayStation (v6.1 and v8B, RaySearch Laboratories AB, Stockholm, Sweden). Four coplanar standard beam angles, including 2 anterior oblique beams with gantry angles of 40^{o} and 320^{o} and 2 posterior oblique beams with gantry angles of 160^{o} and 200^{o} were used for most patients and beam angles were optimized manually such that dose to the parotid glands and swallowing structures was avoided as much as possible. A range shifter of 4.0 cm was used for superficial parts of the targets with a 5-7.5 cm air gap from the patient. The posterior beams were split into range shifter and non-range shifter fields having 6 beams (in 4 angles) in total. For more complex cases, extra beams were used to increase robustness to anatomical changes. Posterior beams were blocked from passing through the shoulders to prevent any range uncertainties due to variations in patient set-up and anterior fields were blocked to avoid irradiation through dental fillings and nasal cavities.

Minimax robust optimization was used to reduce the impact of systematic and random setup errors and systematic range uncertainties on the target coverage and critical organ doses [10]. Until January 2019, we used robustness settings with a setup uncertainty of 5 mm for the IMPT plans, similar to

the margin used in the VMAT plans while creating a CTV-PTV margin. From January 2019 on, the CTV to PTV margin for VMAT plans and set-up uncertainty for the IMPT plans were reduced to 3 mm based on an evaluation of the first cohort of patients treated at our center [11]. The range uncertainty of the IMPT plans was set to ±3.0%. The robust minimum and maximum dose objectives for the CTVs, dose fall-off objectives for the body contour, robust maximum dose objective for serial OARs and dose fall-off, maximum equivalent uniform dose objectives (with dose-volume effect parameter a=1) for the OARs were used during optimization. Finally, plan robustness was evaluated for 28 error scenarios (including 14 translations and ±3% range uncertainties) using the voxel-wise minimum for CTV-coverage and voxel-wise maximum for organ-at risk dose distributions [12].

Step 4: Treatment Plan Comparison

NTCP-profiles of VMAT and IMPT plans were calculated for each individual patient using the NTCP-models of the NIPP. If the difference between VMAT and IMPT corresponded with at least one of the predefined thresholds for ΔNTCP or ΣΔNTCP, then the patient qualified for proton therapy. Otherwise, patients were treated with the original VMAT-plan used for the plan comparison. The model-based selection workflow is depicted in Figure 1.

For patients referred from other centres, the photon plan of the referring centre was used as reference. For the plan comparison, the DICOM-RT data of the referring centre were uploaded to RayStation and these data were used to create the IMPT-plan for the plan comparison. When the plan comparison was positive, patients were actually referred to our proton centre for IMPT preparation and treatment planning.

Statistical Analysis

Descriptive statistics were computed for the group eligible for the model-based approach procedure. The chi-squared test was used to examine the differences between groups by baseline categorical characteristics. For continuous variable comparisons, statistical tests were selected based on type

(related and independent samples) and distribution (normal and non-normal) of the data. The differences in OAR-doses and NTCP-values between proton and photon plans were compared using the Wilcoxon Signed Rank Test or the paired samples T-test, whenever appropriate. Differences between patient groups were compared using Mann–Whitney U-tests or independent samples T-tests, whenever appropriate. For the comparison of >2 groups, the One-Way ANOVA and Kruskal Wallis tests were used for normally and non-normally distributed variables, respectively. All statistical tests were two-sided and a p-value of ≤ 0.05 was considered statistically significant. Analyses were performed using the Statistical Package for Social Sciences (SPSS) for Windows, version 21.0 (SPSS Inc., Chicago, ILL, USA).

Results

The first patient in the Netherlands, selected according to the model-based approach was treated with proton therapy in our center in January 2018. From January 2018 to September 2019, 227 patients with HNC were treated with definitive radiotherapy with or without systemic treatment (chemotherapy or cetuximab). Twenty-seven (12%) of these patients were not considered eligible for proton therapy for various reasons, including synchronous tumor other than HNC region (n=5), rapid progression (n=3), psychologically unsuitable (n=2), metal implants and subsequent CT artefacts (n=3), and/or unsuitable for the immobilization devices requiring a high level of hyperextension (n=14). Twenty-eight (12%) patients were not included into the model-based selection procedure because of tumor locations other than pharynx, larynx or oral cavity.

The remaining 172 (76%) patients went through in the MBS procedure. For all 172 patients, a model-based optimized VMAT plan was created and the NTCP-profile was evaluated. In 31 patients (14%), the PST was negative such that creating an IMPT plan for comparison was deemed unnecessary (Figure 2). These patients were predominantly those with stage I or II tumors, tumors located in the larynx, without baseline complaints and were treated with conventional or accelerated radiotherapy

(Table 2). Even though the PTVs had a considerable overlap with the PCM inferior, supraglottic area and the cricopharyngeal muscle, almost no overlap was observed with other OARs in these patients which reflected the dose distribution in OARs (Figure 3a, Appendix A and B). The median NTCP-values of the VMAT plans were significantly lower in this patient group compared to those qualified for plan comparison (Appendix C).

In the remaining 141 patients (62%), a plan comparison was performed. In these patients, a model-based optimized IMPT plan was generated and used to calculate the ΔNCTP-values between photon and proton plan. The ΔNCTP-thresholds as defined in the NIPP were met in 80 out of 227 patients (35%). Of these, 8 patients could not receive proton therapy because of limited capacity at the proton facility during the ramp-up period. Overall, from January 2018 to May 2019, 155 patients (68%) were treated with photons and 72 patients (32%) were treated with protons. Of these 72 patients, 13 (4 with nasopharyngeal and 9 with oropharyngeal cancer) were referred from other institute and evaluated in model-based selection procedure by comparing NTCP values the photon of the referring center and NCTP values of the proton plan of our centre. They were all qualified for and treated with proton therapy in our institute.

Patients who qualified for proton therapy had more locally advanced disease, pharyngeal tumors and/or baseline complaints and were treated with RT in combination with systemic treatment (Table 2). In addition, the PTVs had more overlap with certain OARs, such as the oral cavity, parotid glands, PCM superior, ipsilateral submandibular gland but less with others, such as the PCM inferior, the cricopharyngeal muscle and supraglottic area (Appendix A). Similarly, the D_{mean} of the oral cavity, PCM superior, parotid glands and ipsilateral submandibular gland were significantly higher in both VMAT and IMPT plans among those who qualified for proton therapy as compared to patients who did not qualify for protons, whereas the D_{mean} of PCM inferior, cricopharyngeal muscle and supraglottic area were higher in patients who did not qualify for protons (Figure 3b; Appendix B).

The mean doses in all OARs were substantially reduced with IMPT compared with VMAT, both in patients who qualified for protons and those who did not. Nevertheless, the ΔDose between VMAT and IMPT plans for the oral cavity, parotid glands, PCM inferior, cricopharyngeal muscle and supraglottic area was significantly higher in patients qualifying for protons compared with those in non-qualifying patients. The ΔDose for the PCM superior, PCM medius, submandibular glands and cervical oesophagus was similar between the two groups (Figure 3b, Appendix B).

The median NTCP-values of the VMAT plans were significantly higher in patients who qualified for proton therapy compared with others. The average NTCP-values obtained with IMPT were significantly lower than those obtained with VMAT. Moreover, these reductions were significantly higher for all endpoints in patients that qualified for protons compared with those observed in non-qualifying patients (Appendix C).

Of the 141 patients qualified for plan comparison, 80 (57%) eventually qualified for protons: 31 (22%), 16 (11%), 25 (18%) and 8 (6%) were selected based on the ΔNTCP value of higher than thresholds for dysphagia, xerostomia, ΣΔNTCP of xerostomia + dysphagia, and TFD, respectively. Thirty patients (21%) met two or more ΔNTCP-criteria. Figure 4 shows the decision tree and number of patients who qualified and did not qualify for proton therapy based on the plan comparison.

Discussion

This is the first report reporting on the clinical implementation of model-based selection for proton therapy in HNC. During the first 20 months, 172 out of 227 patients referred for definitive radiotherapy, with or without systemic treatment, were eligible for the model-based selection procedure, of which 80 patients (35%) eventually qualified for proton therapy.

Patients who were not selected for a plan comparison had mainly early-staged laryngeal cancer and no baseline complaints. These patients had significantly lower risks of xerostomia and dysphagia compared to patients selected for a plan comparison. Patients who eventually qualified for proton therapy had more often locally advanced nasopharyngeal or oropharyngeal cancers with relatively higher dose levels to the nearby OARs such as the parotid glands, oral cavity and PCM superior and were thus at higher risk of developing xerostomia and dysphagia. The primary tumor location also explains the differences in the percentage of OARs overlapping parts with the PTV and differences in dose distributions among patients who eventually qualified for proton therapy and who did not. Most patients were selected based on risk reduction in dysphagia or the summed risk reduction for grade $\geq$ II toxicities. This is in line with the findings of a previous study by Jakobi et al. who reported a higher risk reduction in swallowing-related side effects and higher ΔNTCP-values for patients with tumors in the upper head and neck area [13].

In the Netherlands, full reimbursement for novel treatments is only possible when these are approved by the National Health Care Institute [14]. Approval was obtained for proton therapy in head and neck cancer in November 2017. Consequently, treatment costs in all patients treated with proton therapy were fully reimbursed. However, in a recent study, multivariable analysis evaluating the insurance approval rates and subsequent treatment delays by institutions in USA, showed that coverage was associated with insurance category and not with other factors such as diagnosis, reirradiation, trial enrolment, or the American Society for Radiation Oncology model policy guidelines [15]. Moreover, the rates of initial denial increased significantly from 55% to 74% over a 3-year time-period. Prior authorization resulted in an average treatment delay of 3 weeks (up to 4 months) and led to the termination of radiotherapy in 19% of denied patients [15]. Furthermore, in a study with a larger cohort consisting of lung and HNC patients, the approval rates by insurance companies decreased significantly if there was a plan comparison study demonstrating the benefit of proton therapy for the patient [16]. These results indicate a major variability among countries and stress the need for international guidelines and collaboration between stakeholders on reimbursement to

ensure better health-care for patients with specific types of tumors with a significant expected benefit from proton therapy in terms of toxicity reduction [17].

The model-based approach therefore provides an effective, fair and open tool to legitimize why a patient is selected or not for proton therapy from the patient's, clinician's and payers' point of view. A further advantage is how flexible and dynamic the approach is, since indication protocols can be updated over time based on updated and improved NTCP-models or by adding new NTCP models for other relevant complications [4].

Our first experience showed that model-based selection is time-consuming and requires considerable resources. However, after a short learning period, this approach proved to be logistically feasible and could be streamlined further in terms of improved selection for the plan comparison. The median time between the first consultation and the first fraction of the radiotherapy was 16 days and 14 days in patients who were, and were not, selected for plan comparison, respectively. Moreover, less time was required for the initial plan comparison due to improved expertise and experience of the team involved in treatment planning. We expect that resources for plan comparison can be further reduced by the implementation of automated treatment planning procedures, e.g. using artificial intelligence [18–21].

After lively discussions and a detailed literature search, 3 NTCP models were selected in this initial phase of model-based selection. It should be noted, that these currently used NTCP-models are based on the assessment at 6 months after completion and thus does not cover the full spectrum of acute and late toxicities. In the near future, more comprehensive NTCP-profiles will become available based on multicenter prospective data registration programs as initiated in the Netherlands [22,23].

In conclusion, model-based selection of patients with HNC for proton therapy is clinically feasible. Approximately one third of HNC patients qualify for protons and these patients are the most likely to benefit from protons in terms of prevention of late radiation-induced side effects.

Appendices

Table 1 in Appendix A, Table 1 in Appendix B, Table 1 in Appendix C.

Acknowledgement

This work described in this paper has been granted by the European Union (Project number: 730983: INSPIRE-project).

References

[1] Langendijk JA, Lambin P, De Ruysscher D, 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;107:267–73. https://doi.org/10.1016/j.radonc.2013.05.007.

[2] Landelijk Platform Protonentherapie (LPPT). Landelijk Indicatieprotocol Protonentherapie Hoofdhals 2017. http://www.nvro.nl/publicaties/rapporten.html.

[3] Nation-wide incidence of Head and Neck Cancers 2019. https://www.cijfersoverkanker.nl/selecties/Dataset_2/img5ca32fcf8c644.

[4] CVZ. Protonentherapie bij model-based indicaties (hoofdhalstumoren, mammacarcinoom, longcarcinoom en prostaatcarcinoom) en bij overige indicaties. 304 2011. https://www.zorginstituutnederland.nl/publicaties/standpunten/2011/08/22/protonentherapie-indicaties-voor-protonentherapie-deel-2---model-based-indicaties-hoofd-halstumoren-mammacarcinoom-longcarcinoom-en-prostaatcarcinoom.

[5] Kierkels RGJ, Wopken K, Visser R, Korevaar EW, van der Schaaf A, Bijl HP, et al. Multivariable normal tissue complication probability model-based treatment plan optimization for grade 2–4 dysphagia and tube feeding dependence in head and neck radiotherapy. Radiother Oncol 2016;121:374–80. https://doi.org/10.1016/j.radonc.2016.08.016.

[6] Brouwer CL, Steenbakkers RJHM, Bourhis J, Budach W, Grau C, Grégoire V, et al. CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines. Radiother Oncol 2015. https://doi.org/10.1016/j.radonc.2015.07.041.

[7] Beetz I, Schilstra C, van Luijk P, Christianen ME, Doornaert P, Bijl HP, et al. External validation of three dimensional conformal radiotherapy based NTCP models for patient-rated xerostomia and sticky saliva among patients treated with intensity modulated radiotherapy. Radiother Oncol 2012. https://doi.org/10.1016/j.radonc.2011.11.006.

[8] Christianen MEMC, Schilstra C, Beetz I, Muijs CT, Chouvalova O, Burlage FR, et al. Predictive modelling for swallowing dysfunction after primary (chemo)radiation: Results of a prospective observational study. Radiother Oncol 2012;105:107–14. https://doi.org/10.1016/j.radonc.2011.08.009.

[9] Wopken K, Bijl HP, Van Der Schaaf A, Van Der Laan HP, Chouvalova O, Steenbakkers RJHM, 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;113:95–101. https://doi.org/10.1016/j.radonc.2014.09.013.

[10] Fredriksson A, Forsgren A, Hårdemark B. Minimax optimization for handling range and setup uncertainties in proton therapy. Med Phys 2011. https://doi.org/10.1118/1.3556559.

[11] Wagenaar D, Kierkels RGJ, van der Schaaf A, Meijers A, Scandurra D, Sijtsema M, et al. Robustness Evaluation Using Dose Accumulation in Head and Neck IMPT. Int J Radiat Oncol 2019. https://doi.org/10.1016/j.ijrobp.2019.06.854.

[12] Korevaar EW, Habraken SJM, Scandurra D, Kierkels RGJ, Unipan M, Eenink MGC, et al. Practical robustness evaluation in radiotherapy – A photon and proton-proof alternative to PTV-based plan evaluation. Radiother Oncol 2019. https://doi.org/10.1016/j.radonc.2019.08.005.

[13] Jakobi A, Bandurska-Luque A, Stützer K, Haase R, Löck S, Wack LJ, 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. https://doi.org/10.1016/j.ijrobp.2015.04.031.

[14] Beoordeling "stand van de wetenschap en praktijk" 2015. https://www.zorginstituutnederland.nl/publicaties/rapport/2015/01/15/beoordeling-stand-van-de-wetenschap-en-praktijk.

[15] Gupta A, Khan AJ, Goyal S, Millevoi R, Elsebai N, Jabbour SK, et al. Insurance Approval for Proton Beam Therapy and its Impact on Delays in Treatment. Int J Radiat Oncol Biol Phys 2019. https://doi.org/10.1016/j.ijrobp.2018.12.021.

[16] Ning MS, Gomez DR, Shah AK, Kim CR, Palmer MB, Thaker NG, et al. The Insurance Approval Process for Proton Radiation Therapy: A Significant Barrier to Patient Care. Int J Radiat Oncol Biol Phys 2019. https://doi.org/10.1016/j.ijrobp.2018.12.019.

[17] Shah A, Ricci KI, Efstathiou JA. Beyond a moonshot: insurance coverage for proton therapy. Lancet Oncol 2016;17:559–61. https://doi.org/10.1016/s1470-2045(16)00171-6.

[18] Delaney AR, Dahele M, Tol JP, Kuijper IT, Slotman BJ, Verbakel WFAR. Using a knowledge-based planning solution to select patients for proton therapy. Radiother Oncol 2017. https://doi.org/10.1016/j.radonc.2017.03.020.

[19] Kierkels RGJ, Fredriksson A, Both S, Langendijk JA, Scandurra D, Korevaar EW. Automated Robust Proton Planning Using Dose-Volume Histogram-Based Mimicking of the Photon Reference Dose and Reducing Organ at Risk Dose Optimization. Int J Radiat Oncol Biol Phys 2019. https://doi.org/10.1016/j.ijrobp.2018.08.023.

[20] Delaney AR, Dong L, Mascia A, Zou W, Zhang Y, Yin L, et al. Automated knowledge-based intensity-modulated proton planning: An international multicenter benchmarking study. Cancers (Basel) 2018. https://doi.org/10.3390/cancers10110420.

[21] Walsh S, de Jong EEC, van Timmeren JE, Ibrahim A, Compter I, Peerlings J, Sanduleanu S, Refaee T, Keek S, Larue RTHM, van Wijk Y, Even AJG, Jochems A, Barakat MS, Leijenaar RTH1

LP. Decision Support Systems in Oncology. JCO Clin Cancer Inform 2019;3:1–9. https://doi.org/10.1200/CCI.18.00001.

[22] Langendijk JA, Orecchia R, Haustermans K, Zips D, Balosso J, Lacombe D, et al. Prospective data registration and clinical trials for particle therapy in Europe. Radiother Oncol 2018;128:9–13. https://doi.org/10.1016/j.radonc.2018.06.001.

[23] Langendijk JA, Boersma LJ, Rasch CRN, van Vulpen M, Reitsma JB, van der Schaaf A, et al. Clinical Trial Strategies to Compare Protons With Photons. Semin Radiat Oncol 2018. https://doi.org/10.1016/j.semradonc.2017.11.008.

Figure Captions

Figure 1. Model-based selection workflow.

Figure 2. A selection flowchart of patients with HNC for photon or proton therapy. *Tumor locations other than pharynx, larynx or oral cavity.

Figure 3a. The average D_{mean} values of organ at risks in VMAT plans of the 141 patients qualifying (red area) and 31 patients non-qualifying (purple area) for a plan comparison.

Figure 3b. The average D_{mean} values of organ at risks of the 80 patients qualifying and 60 patients non-qualifying for proton therapy. The mean doses for VMAT and IMPT plans are shown in red and green area, respectively.

Figure 4. For patients who qualified for plan comparison study (n=141), the decision tree and number of patients meeting the minimum ΔNTCP criteria for proton therapy.

Highlights

- Model-based selection of patients with HNC for proton is clinically feasible**.**
- Around 35% of HNC patients qualify for protons.
- Most patients are selected based on ΔNTCP in dysphagia-related models.
- Patients with advanced and pharyngeal tumor have higher probability to be selected.

Appendix A

Table 1. The percentages of overlapping parts of organs at risks with PTV54.25+5 mm (OAR∩(PTV54.5 + 5 mm)) in patient groups.

	Patients NOT qualified for plan comparison	**Patients qualified for plan comparison**		
		NOT qualified for protons	**Qualified for protons**	**P value**
	Overlapping parts of OARs in the models with PTV (%)			
Oral Cavity	0.0% (0.0-0.0)	8.8% (0.2-42.4)	32.8% (16.6-51.3)	<0.001
Contralateral Parotid	0.0% (0.0-0.0)	15.7% (8.1-20.7)	20.1% (10.9-27.7)	<0.001
PCM Superior	0.0% (0.0-0.2)	52.5% (25.3-79.0)	76.0% (57.2-90.2)	<0.001

PCM Inferior	72% (42.1-83.8)	79.5% (35.4-99.6)	37.9% (23.0-71.3)	<0.001
Cricopharyngeus	30.8% (14.1-40.5)	38.1% (11.4-81.7)	15.6% (5.5-34.5)	0.002
Overlapping parts of other OARs with PTV (%)				
Ipsilateral Parotid	0.0% (0.0-0.0)	24.3% (15.7-35.8)	34.6% (24.6-49.9)	<0.001
Ipsilateral Submandibular	0.0% (0.0-1.2)	100.0% (50.4-100.0)	100.0% (100.0-100.0)	<0.001
Contralateral Submandibular	0.0% (0.0-0.0)	48.7% (34.5-95.2)	53.0% (40.2-78.7)	<0.001
PCM Medius	5.3% (0-17.7)	71.9% (45.3-95.5)	73.8% (57.5-92.4)	<0.001
Supraglottis	59.9% (49.7-82)	79.3% (44.9-98.3)	42.0% (24.8-86.7)	0.003
Cervical Esophagus	0.0% (0.0-0.0)	3.0% (0.0-38.8)	0.6% (0.0-16.5)	<0.001

Median(Q25-Q75) values are given.

Appendix B

Table 1. The comparison of VMAT vs. IMPT plans and patient groups in terms of D_{mean} of the OARs.

		Patients NOT qualified for plan comparison	**Patients qualified for plan comparison**		**P value** *(group comparison)*
			NOT qualified for protons	**Qualified for protons**	
		D_{mean} of the OARs in the NTCP models (Gy)			
Oral Cavity	VMAT	0.5 (0.3-0.7)	26.9 (17-40.7)	41.1 (34.2-47.8)	<0.001
	IMPT		9.6 (1.6-32.3)	24.2 (16.9-35.5)	<0.001
	Δdose		10.9 (8.3-15.0)	14.6 (10.6-17.5)	0.018
	P value *(VMAT vs. IMPT)*		<0.001	<0.001	
Contralateral	VMAT	0.2 (0.1-0.4)	14.9 (11.8-19.9)	20.6 (16.5-24.2)	<0.001

Parotid	IMPT		12.8 (9.1-16.8)	14.2 (10.9-18.5)	0.099
	Δdose		2.5 (1.5-4.2)	6.3 (3.1-8.0)	<0.001
	P value *(VMAT vs. IMPT)*		<0.001	<0.001	
PCM Superior	VMAT	0.5 (0.2-0.8)	37.9 (28.5-55.8)	55.8 (48.1-61.5)	<0.001
	IMPT		30.1 (17.9-53.5)	51.9 (40.2-56.5)	0.001
	Δdose		6.2 (1.9-12.5)	5.2 (2.7-8.6)	0.705
	P value *(VMAT vs. IMPT)*		<0.001	<0.001	
PCM Inferior	VMAT	49.8 (41.9-56.0)	54.4 (34.6-64.0)	38.1 (29.9-50.8)	0.003
	IMPT		48.0 (26.2-62.0)	23.9 (20.3-44.9)	<0.001
	Δdose		3.0 (0.6-7.5)	7.7 (3.8-11.7)	<0.001
	P value *(VMAT vs. IMPT)*		<0.001	<0.001	
Cricopharyngeus	VMAT	31.1 (24.1-39.9)	39.8 (25.7-52.3)	30.5 (23.7-39.3)	0.010
	IMPT		30.3 (16.9-50.7)	19.2 (14.2-28.7)	<0.001
	Δdose		4.3 (0.7-9.2)	9.6 (4.6-13.4)	<0.001
	P value *(VMAT vs. IMPT)*		<0.001	<0.001	
		D_{mean} of the other OARs (Gy)			
Ipsilateral Parotid	VMAT	0.3 (0.2-0.5)	20.6 (16.0-28.6)	28.5 (23.4-36.8)	<0.001
	IMPT		19.2 (14.0-25.3)	23.7 (19.4-34.4)	0.005
	Δdose		2.3 (1.0-3.5)	4.2 (1.8-5.9)	<0.001
	P value *(VMAT vs. IMPT)*		<0.001	<0.001	
Ipsilateral Submandibular	VMAT	3.6 (1.4-11.3)	60.7 (41.7-65.8)	64.5 (58.5-67.3)	<0.001
	IMPT		58.8 (39.5-66.0)	63.8 (58.4-66.3)	0.019
	Δdose		0.7 (0.0-2.4)	0.7 (0.2-1.6)	0.835
	P value *(VMAT vs. IMPT)*		<0.001	<0.001	
Contralateral Submandibular	VMAT	2.3 (1.2-4.8)	42.3 (35.9-53.5)	44.9 (38.1-52.3)	<0.001
	IMPT		38.9 (29.9-53.6)	38.3 (32.6-51.0)	0.711
	Δdose		2.6 (0.0-7.9)	2.9 (0.4-7.7)	0.615
	P value *(VMAT vs. IMPT)*		<0.001	<0.001	
PCM Medius	VMAT	13.8 (4.2-21.8)	51.3 (39.7-60.3)	51.4 (44.3-61.5)	<0.001
	IMPT		46.0 (35.4-58.8)	46.0 (38.3-60.3)	0.957
	Δdose		3.3 (1.1-7.1)	3.4 (0.9-7.1)	0.597
	P value *(VMAT vs. IMPT)*		<0.001	<0.001	
Supraglottis	VMAT	41.3 (33.9-53.1)	55.7 (44.2-67.3)	44.7 (36.8-61.2)	0.002
	IMPT		55.0 (38.4-66.1)	35.6 (25.2-58.4)	0.001
	Δdose		1.9 (0.3-5.6)	5.4 (1.3-12.4)	<0.001
	P value *(VMAT vs. IMPT)*		<0.001	<0.001	
Cervical Esophagus	VMAT	2.7 (1.9-4.0)	25.9 (19.9-40.6)	25.7 (18.7-37.7)	<0.001
	IMPT		15.9 (11.1-37.0)	15.3 (8.5-26.6)	0.650
	Δdose		5.4 (1.9-12.6)	8.0 (4.3-13.5)	0.302
	P value *(VMAT vs. IMPT)*		<0.001	<0.001	

Median(Q25-Q75) values are given.

Appendix C

Table 1. The comparison of VMAT vs. IMPT plans and patient groups in terms of NTCP values.

		Patients NOT qualified for plan comparison	**Patients qualified for plan comparison**		**P value** *(group comparison)*
			NOT qualified for protons	**Qualified for protons**	
Xerostomia	VMAT	18.3% (18.2-21.3)	38.4% (32.0-45.0)	44.4% (37.2-56.2)	0.001
	IMPT		34.9% (29.7-43.5)	36.7% (30.2-50.0)	0.140
	ΔNTCP		2.9% (1.8-5.0)	7.7% (3.7-9.7)	< 0.001
	P value (*VMAT vs. IMPT*)		<0.001	<0.001	
Dysphagia	VMAT	3.6% (3.6-3.7)	20.3% (10.3-26.3)	30.2% (23.9-49.1)	< 0.001

	IMPT		13.3% (6.0-21.8)	20.1% (14.9-38.5)	< 0.001
	ΔNTCP		4.6% (3.4-6.5)	8.7% (6.7-10.4)	< 0.001
	P value (*VMAT vs. IMPT*)		<0.001	<0.001	
ΣNTCP of xerostomia + dysphagia	VMAT	22.0% (21.8-29.5)	56.1% (48.4-74.7)	78.5% (64.9-98.5)	< 0.001
	IMPT		48.0% (38.9-70.7)	62.0% (48.3-82.9)	0.001
	ΣΔNTCP		7.9% (5.0-10.6)	15.5% (13.9-16.9)	< 0.001
	P value (*VMAT vs. IMPT*)		<0.001	< 0.001	
Tube feeding dependence	VMAT	0.3% (0.3-0.4)	4.5% (2.3-7.5)	7.1% (3.5-13.4)	0.001
	IMPT		3.3% (1.4-5.4)	4.5% (2.1-9.4)	0.020
	ΔNTCP		0.9% (0.5-1.9)	2.3% (1.2-4.6)	< 0.001
	P value (*VMAT vs. IMPT*)		<0.001	<0.001	

Median(Q25-Q75) values are given.

Tables

Table 1. The regression coefficients of NTCP models after external validation for determining NTCP profile and related ΔNTCP thresholds accepted in the first version of National Indication Protocol Proton therapy (NIPP) for patient selection for proton therapy in head and neck cancer. The lowest panel indicates the ΔNTCP threshold for the three toxicities.

Predictors	**Patient-rated grade ≥ 2 xerostomia**	**Physician-rated grade ≥ 2 dysphagia**	**Tube feeding dependence**
Constant (B0)	-1.507	-3.303	-6.849
D_{mean} Contralateral parotid	0.052		0.022
D_{mean} Oral cavity		0.024	
D_{mean} PCM Superior		0.024	0.030
D_{mean} PCM Inferior			0.013
D_{mean} Cricopharyngeus			0.008
Baseline xerostomia: Moderate (EORTC QLQ-H&N35 - Q40: score 3)	0.525		

Severe (EORTC QLQ-H&N35 - Q40: score 4)	1.482		
Baseline Grade II-III dysphagia		0.967	
Treatment modality:			
Accelerated radiotherapy			0.198
Concurrent chemoradiation			1.101
Radiotherapy + cetuximab			1.716
Weight loss prior to treatment: 1-10%			0.317
>10%			1.178
T-stage: stage III-IV			0.680
ΔNTCP threshold	**≥ 10%**	**≥ 10%**	**≥ 5%**
	ΣΔNTCP ≥ 15%		

Table 2. Baseline characteristics of patients qualifying and not qualifying for proton therapy.

		Qualified for plan comparison				
		No (n=31)	Yes (n=141)			
			Qualified for protons		Total	
Baseline characteristics			No (n=61)	Yes (n=80)	(n=172)	P value
T stage	T1-T2	24 (77%)	18 (30%)	30 (38%)	72 (42%)	<0.001
	T3-T4	7 (23%)	43 (70%)	50 (63%)	100 (58%)	
N stage	N0	30 (97%)	35 (57%)	33 (41%)	70 (41%)	<0.001
	N+	1 (3%)	26 (43%)	47 (59%)	102 (59%)	
Stage	Stage I	15 (48%)	1 (2%)	2 (3%)	18 (10%)	<0.001
	Stage II	8 (26%)	4 (7%)	1 (1%)	13 (8%)	
	Stage III	4 (13%)	23 (38%)	19 (24%)	46 (27%)	
	Stage IVA	4 (13%)	31 (51%)	52 (65%)	87 (51%)	
	Stage IVB	0 (0%)	2 (3%)	6 (8%)	8 (5%)	
Tumor location	Oropharynx	0 (0%)	28 (46%)	51 (64%)	79 (46%)	<0.001
	Larynx	29 (94%)	29 (48%)	8 (10%)	66 (38%)	
	Nasopharynx	0 (0%)	0 (0%)	10 (13%)	10 (6%)	
	Hypopharynx	0 (0%)	4 (7%)	6 (8%)	10 (6%)	
	Oral cavity	2 (6%)	0 (0%)	5 (6%)	7 (4%)	

Baseline xerostomia	None	28 (90%)	36 (59%)	42 (53%)	106 (62%)	0.031
	A little	2 (6%)	19 (31%)	28 (35%)	49 (28%)	
	Quite a bit	1 (3%)	6 (10%)	10 (13%)	17 (10%)	
Baseline dysphagia	Grade 0-1	31 (100%)	47 (77%)	49 (61%)	127 (74%)	<0.001
	Grade 2-3	0 (0%)	14 (23%)	31 (39%)	45 (26%)	
Baseline weight loss	None	25 (80%)	39 (64%)	43 (54%)	107 (62%)	0.003
	1-10%	3 (10%)	19 (31%)	23 (29%)	45 (26%)	
	> 10%	3 (10%)	3 (5%)	14 (18%)	20 (12%)	
Treatment modality	Chemoradiation	1 (3%)	19 (31%)	36 (45%)	56 (33%)	0.001
	Conventional RT	19 (61%)	25 (41%)	25 (31%)	69 (40%)	
	Accelerated RT	11 (35%)	14 (23%)	13 (16%)	38 (22%)	
	RT+cetuximab	0 (0%)	3 (5%)	6 (8%)	9 (5%)	
Time*		14 (12-18)	16 (14-20)		15 (14-19)	0.003

*Median (Q25-Q75) interval (day) between first consultation and the first fraction of radiotherapy.