University of Groningen

Quality of life and toxicity guided treatment plan optimisation for head and neck cancer

van der Laan, Hans Paul; van der Schaaf, Arjen; Van den Bosch, Lisa; Korevaar, Erik W; Steenbakkers, Roel J H M; Both, Stefan; Langendijk, Johannes A

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Original Article

Quality of life and toxicity guided treatment plan optimisation for head and neck cancer ☆

Hans Paul van der Laan *, Arjen van der Schaaf, Lisa Van den Bosch, Erik W. Korevaar, Roel J.H.M. Steenbakkers, Stefan Both, Johannes A. Langendijk

Department of Radiation Oncology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

ABSTRACT

Purpose: To evaluate the feasibility of semi-automatic Quality of Life (QOL)-weighted normal tissue complication probability (NTCP)-guided VMAT treatment plan optimisation in head and neck cancer (HNC) and compare predicted QOL to that obtained with conventional treatment.
Materials and methods: This study included 30 HNC patients who were treated with definitive radiotherapy. QOL-weighted NTCP-guided VMAT plans were optimised directly on 80 multivariable NTCP models of 20 common toxicities and symptoms on 4 different time points (6, 12, 18 and 24 months after radiotherapy) and each NTCP model was weighted relative to its impact on QOL. Planning results, NTCP and predicted QOL were compared with the clinical conventional VMAT plans.
Results: QOL-weighted NTCP-guided VMAT plans were clinically acceptable, had target coverage equally adequate as the clinical plans, but prioritised sparing of organs at risk (OAR) related to toxicities and symptoms that had the highest impact on QOL. NTCP was reduced for, e.g., dysphagia (−6.1% for ≥grade 2/−7.6% for ≥grade 3) and moderate-to-severe fatigue/speech problems/hoarseness (−0.7%/−1.5%/−2.5%) at 6 months, respectively. Concurrently, the average NTCP of toxicities related to salivary function increased with +0.4% to +5.7%. QOL-weighted NTCP-guided plans were produced in less time, were less dependent on the treatment planner experience and yielded more consistent results. The average predicted QOL improved by 0.7, 0.9, 1.0, and 1.1 points on a 0–100 scale (p < 0.001) at 6, 12, 18, and 24 months, respectively, compared to the clinical plans.
Conclusion: Semi-automatic QOL-weighted NTCP-guided VMAT treatment plan optimisation is feasible. It prioritised sparing of OARs related to high-impact toxicities and symptoms and resulted in a systematic improvement of predicted QOL compared to conventional VMAT.

The introduction of intensity modulated radiation therapy (IMRT) introduced a paradigm shift in radiation oncology. With this revolutionary new method, planning target volumes (PTV) and organs at risk (OAR) are assigned objectives and weights so that a treatment planning system can produce the most optimal plan [1]. With IMRT and volumetric modulated arc therapy (VMAT), treatment plan optimisation includes a series of iterations and manual adjustments of the objective values and weights until a clinically acceptable solution is reached [2]. In recent years, this process has become exponentially complex as an increasing

number of OARs have become part of the equation. In clinical practise, it is often challenging to find the optimal balance in sparing the multitude of OARs in a way that is beneficial to patients.

Multivariable normal tissue complication probability (NTCP) models can aid in finding this optimal balance, as an OAR that most affects the NTCP, can be spared with priority. In the case of head and neck cancer (HNC), recent publications have produced a comprehensive list of NTCP models [3,4]. Because the number of toxicities for which NTCP models are available has also significantly increased, it has become even more complex to find the optimal solution, especially when optimisation is performed iteratively and manually. NTCP-guided treatment plan optimisation has been proposed previously to solve this problem. Kierkels et al. used NTCP models for a limited number of equally weighted toxicities as optimisation functions in the treatment planning system with IMRT and demonstrated this method to be feasible, efficient and to result in clinically realistic treatment plans [5]. However, with the body of NTCP models currently available it remains unclear

which toxicities should receive priority in being prevented, i.e., which toxicities should receive higher weights in the optimisation process, in order to achieve the optimal treatment plan in terms of highest QOL [6].

We recently reported on a QOL model that can be used to prioritise the prevention of high-impact toxicities and symptoms in the context of a comprehensive priority-weighted toxicity profile for HNC patients [7]. This QOL model can be combined with the NTCP models for the various toxicities and symptoms [3]. The combined models provide for: (1) optimisation functions for semi-automatic QOL-weighted NTCP-guided treatment plan optimisation; and (2) QOL-guided treatment plan comparison based on dose distributions, considering patient baseline and treatment characteristics. In this way, the plan with the highest expected QOL can be selected as the clinical plan for each patient.

The purpose of this study was to test the feasibility of semi-automatic QOL-weighted NTCP-guided treatment optimisation for VMAT and to compare dose in OARs, NTCP and QOL with QOL-weighted NTCP-guided VMAT to that of existing clinical plans based on conventional OAR-based VMAT treatment plan optimisation.

Materials and methods

Patients

The population of this study was composed of 30 patients who had previously undergone definitive radiotherapy for HNC and required elective or therapeutic treatment of the neck. The patients were selected to represent our patient population (Supplementary Table 1). All patient data was obtained as part of a prospective data registration programme within the framework of routine clinical practise (clinicaltrial.gov NCT02435576) and the programme was reviewed and exempted from the ethical approval requirement by the hospital ethics committee. Patients were given the opportunity to withdraw their consent for their data to be used for research purposes at any time and all data was pseudonymised before use.

Target volumes and organs at risk

Contrast enhanced planning CT scans were acquired in treatment position with a slice thickness and index of 2 mm. Target volumes, critical structures and OARs were defined according to

international consensus guidelines [8,9]. Planning target volumes (PTV) were restricted to 5 mm within the skin surface for the purpose of dose evaluation. Dose was delivered to the PTV following a simultaneous integrated boost (SIB) schedule of 35 fractions, with a daily dose of 1.55 Gy prescribed to elective nodal regions (PTV54) and a daily dose of 2.00 Gy prescribed to a high risk PTV (PTV70). Critical structures included the spinal cord, brainstem, optic nerves, and optical chiasm. OARs related to various common acute and late toxicities and symptoms were included (Tables 1 and 2). These OARs were described previously to be part of our comprehensive NTCP profile [3].

Treatment planning

VMAT treatment planning was performed in the RayStation planning system (RaySearch Laboratories, Stockholm, Sweden). The clinical plans, created in the clinical version, were imported in a research version of the system (version 10B -R, build 10.1.100.0) offering functionalities not yet implemented in the clinical version. Subsequently, QOL-weighted NTCP-guided plans were created. Two full arcs were used with all plans and the primary objectives for all plans were identical: at least 98% of each PTV had to be covered with 95% of the prescribed dose, the maximum doses delivered to the spinal cord, brainstem, optic nerves, and optic chiasm were not allowed to exceed 54 Gy, 60 Gy, 54 Gy and 54 Gy, respectively. The maximum plan dose was not allowed to exceed 77 Gy and the volume receiving 75 Gy was not allowed to be larger than 2 cm^3 (Supplementary Table 2). For all plans, dose optimisation included a series of optimisation sequences, each consisting of 80 automated iterations, and a trial-and-error adaptive adjustment of the objectives' values and weights until a clinical acceptable solution was reached. The Collapsed Cone v5.3 algorithm was used for all final dose calculations. The only difference between clinical and QOL-weighted NTCP-guided plans was how OAR dose was optimised.

OAR dose optimisation

In the clinical conventional plans, the values and weights of the equivalent uniform dose (EUD)-based [10] OAR objectives were balanced with the physical dose-based objectives for the PTVs, critical structures, and general planning objectives. The order of priority in which OARs were spared (Supplementary Table 2), was based on the effect size and frequency of occurrence of the OARs in the

Table 1
Dose in organs at risk.

	Conventional mean ± st. error	QOLO mean ± st. error	QOLO—Conventional mean ± st. error	*p*-value
Organs at risk				
Oral cavity Dmean (Gy)	47.0 ± 2.2	42.2 ± 2.7	−4.8 ± 0.9	<0.001
PCM superior Dmean (Gy)	54.1 ± 2.0	52.9 ± 2.0	−1.2 ± 0.4	0.002
PCM medius Dmean (Gy)	57.4 ± 1.8	54.1 ± 2.2	−3.2 ± 0.7	<0.001
PCM inferior Dmean (Gy)	45.8 ± 2.7	39.3 ± 3.1	−6.6 ± 1.2	<0.001
Buccal mucosa Dmean (Gy)	39.7 ± 2.1	35.8 ± 2.5	−4.0 ± 0.7	<0.001
Mandible Dmean (Gy)	40.3 ± 1.5	39.3 ± 1.6	−1.0 ± 0.3	0.005
Supraglottic larynx Dmean (Gy)	53.8 ± 2.6	47.6 ± 3.3	−6.1 ± 1.0	<0.001
Glottic larynx Dmean (Gy)	43.7 ± 2.8	38.5 ± 3.2	−5.2 ± 1.2	<0.001
Arytenoid IL Dmean (Gy)	47.5 ± 3.0	39.5 ± 3.5	−7.9 ± 1.3	<0.001
Arytenoid CL Dmean (Gy)	41.9 ± 3.0	33.7 ± 3.4	−8.2 ± 1.2	<0.001
Parotid IL Dmean (Gy)	33.6 ± 1.8	37.1 ± 1.6	3.5 ± 0.6	<0.001
Parotid CL Dmean (Gy)	23.3 ± 1.7	27.4 ± 1.8	4.1 ± 0.7	<0.001
Submandibular IL Dmean (Gy)	64.4 ± 1.0	63.7 ± 1.0	−0.7 ± 0.3	0.018
Submandibular CL Dmean (Gy)	52.8 ± 2.2	52.4 ± 2.1	−0.4 ± 0.4	0.393
CT-scanned patient Dmean (Gy)	13.0 ± 0.6	12.3 ± 0.6	−0.7 ± 0.2	<0.001
Integral dose (Gy * m^3)	0.162 ± 0.007	0.152 ± 0.006	−0.009 ± 0.002	<0.001

Abbreviations: Conventional = Clinical plans optimised by using equivalent uniform dose OAR based objectives and weights; IL/CL = Ipsilateral/Contralateral; PCM = pharyngeal constrictor muscle; QOLO = Quality of Life-optimised.

Hans Paul van der Laan, A. van der Schaaf, L. Van den Bosch et al.

Radiotherapy and Oncology 162 (2021) 85–90

Table 2
Normal tissue complication probabilities.

	6 months				12 months*	18 months*	24 months*
	Conventional	QOLO	QOLO—Conventional	p-value	QOLO—Conventional	QOLO—Conventional	QOLO—Conventional
	% ± st. error	% ± st. error	% point ± st. error		% point ± st. error	% point ± st. error	% point ± st. error
Swallowing							
Dysphagia, grade 2–4	53.7 ± 4.8	47.5 ± 4.8	−6.1 ± 0.7	<0.001	−6.2 ± 0.7	−5.2 ± 0.7	−5.2 ± 0.7
Dysphagia, grade 3–4	27.1 ± 4.6	19.5 ± 3.9	−7.6 ± 1.2	<0.001	−5.7 ± 1.0	−4.9 ± 0.8	−4.9 ± 0.8
Aspiration, grade 2–4	18.9 ± 2.2	17.2 ± 2.0	−1.7 ± 0.3	<0.001	−1.9 ± 0.3	−2.7 ± 0.5	−3.1 ± 0.5
Aspiration, moderate-severe	15.3 ± 2.3	14.4 ± 2.2	−0.9 ± 0.2	<0.001	−0.9 ± 0.2	−1.1 ± 0.2	−1.1 ± 0.2
Salivary							
Xerostomia, moderate-severe	52.9 ± 2.6	54.4 ± 2.6	1.5 ± 0.3	<0.001	1.5 ± 0.3	1.5 ± 0.3	1.5 ± 0.3
Xerostomia, severe	16.7 ± 1.6	17.5 ± 1.7	0.9 ± 0.2	<0.001	0.9 ± 0.2	0.9 ± 0.2	0.7 ± 0.1
Sticky saliva, moderate-severe	39.0 ± 1.9	39.9 ± 1.9	0.9 ± 0.2	<0.001	0.8 ± 0.1	0.6 ± 0.1	0.6 ± 0.1
Sticky saliva, severe	13.7 ± 0.9	14.1 ± 1.0	0.4 ± 0.1	<0.001	0.3 ± 0.1	0.3 ± 0.1	0.4 ± 0.1
Xerostomia, grade 2–4	40.7 ± 2.8	46.4 ± 2.9	5.7 ± 0.9	<0.001	5.5 ± 0.8	5.2 ± 0.8	5.5 ± 0.8
Sticky saliva, grade 2–4	13.6 ± 0.7	13.4 ± 0.7	−0.1 ± 0.2	0.34	−0.1 ± 0.2	−0.1 ± 0.1	−0.1 ± 0.2
Loss of taste, moderate-severe	33.4 ± 0.9	32.7 ± 1.0	−0.7 ± 0.4	0.064	−0.5 ± 0.3	−0.5 ± 0.3	−0.5 ± 0.3
Loss of taste, grade 2–4	26.6 ± 0.9	27.7 ± 0.9	1.1 ± 0.2	<0.001	1.7 ± 0.3	1.3 ± 0.2	1.9 ± 0.4
Speech							
Hoarseness, moderate-severe	14.5 ± 2.3	12.1 ± 2.3	−2.5 ± 0.5	<0.001	−2.0 ± 0.4	−2.0 ± 0.4	−1.8 ± 0.3
Speech problems, moderate-severe	23.7 ± 3.1	22.2 ± 3.0	−1.5 ± 0.2	<0.001	−1.7 ± 0.2	−3.2 ± 0.3	−3.2 ± 0.3
Pain							
Oral pain, moderate-severe	29.8 ± 2.5	28.2 ± 2.5	−1.7 ± 0.3	<0.001	−1.4 ± 0.2	−1.4 ± 0.2	−2.0 ± 0.3
Throat pain, moderate-severe	26.0 ± 2.3	23.4 ± 2.4	−2.6 ± 0.4	<0.001	−2.2 ± 0.4	−2.2 ± 0.4	−2.2 ± 0.4
Jaw pain, moderate-severe	18.5 ± 1.6	18.1 ± 1.6	−0.5 ± 0.2	0.007	−0.5 ± 0.2	−0.4 ± 0.1	−0.4 ± 0.1
General							
Weightloss > 10% over baseline	21.9 ± 2.0	19.4 ± 1.9	−2.5 ± 0.4	<0.001	−2.3 ± 0.4	−2.3 ± 0.4	−1.7 ± 0.2
Nausea and vomiting, moderate-severe	7.6 ± 0.4	7.2 ± 0.4	−0.5 ± 0.1	<0.001	−0.5 ± 0.1	−1.1 ± 0.2	−1.0 ± 0.2
Fatigue, moderate-severe	31.0 ± 2.9	30.3 ± 2.8	−0.7 ± 0.2	<0.001	−0.7 ± 0.2	−0.8 ± 0.2	−0.8 ± 0.2

*see supplementary table for full results at 12,18 and 24 months.
Abbreviations: OAR = Organs at risk; Conventional = Clinical plans optimised by using OAR-based objectives and weights; QOLO = Quality of Life-optimised (QOL-weighted NTCP-guided). Symptoms and toxicities include patient reported symptoms (moderate-severe, severe), physician rated toxicities (grade 2–4, grade 3–4) and weightloss.

comprehensive NTCP profile [3], and not necessarily on the importance of the various toxicities in relation to QOL (Supplementary Table 3). In the QOL-weighted NTCP-guided plans, 80 NTCP-guided optimisation functions (including models for 20 toxicities and symptoms at 4 time points after radiotherapy) were used to optimise the dose in the OAR (Supplementary Appendix) [3,5]. Each of the NTCP-guided optimisation functions was weighted according to the impact of the corresponding toxicity or symptom on QOL (Supplementary Table 3) [7]. For example, the NTCP-guided optimisation function of dysphagia grade 2–4 had a weight of 4%, whereas moderate-severe xerostomia had a weight of 1%. The method of optimising the plans in a trial-and-error adaptive fashion was similar to the conventional clinical plans, except that the NTCP-guided optimisation functions and their weights were never changed during the whole procedure and adjustments were only made to the physical dose-based objectives (PTV, critical structures and general objectives) (Fig. 1). The number of optimisation sequences and treatment time required for QOL-weighted NTCP-guided plans were recorded and compared to the average time needed for conventional VMAT planning in the clinic.

Evaluation of dose plans

Target coverage was determined for all plans by evaluating D98 (the lowest dose in the 98% volume receiving the highest dose) for PTV54 and PTV70. General planning objectives were checked, and the location and size of hotspots, cold spots and other clinically relevant aspects of the dose distribution were visually inspected by experienced dosimetrists to be acceptable for the clinical and QOL-weighted NTCP-guided plans. Homogeneity indices (D2/D98) and conformity indices (volume covered by the 95% isodose line/the volume inside the PTV covered by the 95% isodose line)

were determined for PTV54 and PTV70. Dose-volume parameters were obtained for all relevant structures from corresponding dose-volume histograms (DVH).

By combining baseline characteristics and dose-volume parameters, NTCP values could be calculated. We previously reported on a comprehensive NTCP profile, describing multivariable NTCP models for various endpoints and toxicity domains [3]. We used these models in the current study to calculate NTCP values for each plan and for each of the 20 toxicities at 4 time points. Subsequently, the predicted QOL was calculated for each plan by combining our NTCP and QOL models [3,7] (Fig. 1). A detailed example of this method can be found in the (Supplementary Appendix). In brief, the QOL model enables the calculation of QOL given that the toxicity and symptom outcomes are known, i.e., by multiplying the regression coefficient of each subsequent toxicity or symptom with 0 in the case of a non-event and with 1 in the case of an event. However, as the toxicity and symptom outcomes were not known at the time of treatment planning, NTCP values (ranging from 0 to 1), different for each plan, were used instead of 0 and 1 to predict QOL. QOL was also calculated for each patient for an imaginary plan with a dose of 0 Gy in all OARs, i.e., the QOL we expected the patient to have when zero dose was given to the healthy tissues. This allowed us to calculate a maximum achievable QOL gain as a reference value. Average differences in outcomes between the treatment optimisation methods were tested for statistical significance ($p \leq 0.05$) with a two-tailed Wilcoxon signed-rank test or Student t-test, whenever appropriate.

Results

On average, conventional VMAT plans were created in 3 hours and required plan evaluation and manual adjustments at each step.
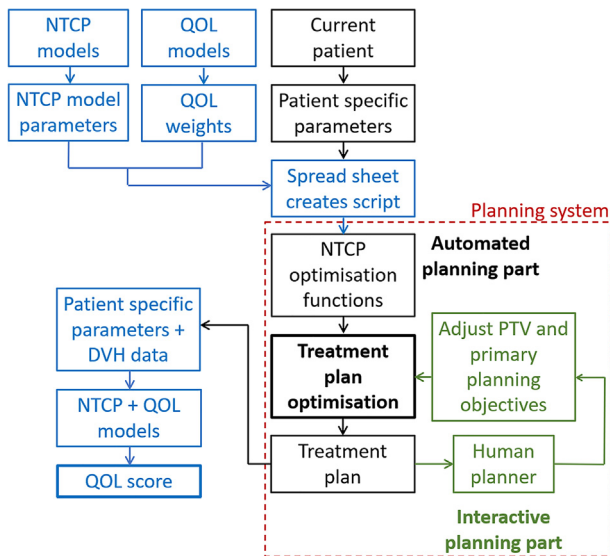
Fig. 1. Flowchart for treatment planning script creation, treatment planning, and treatment plan evaluation. Normal tissue complication probability (NTCP) models, Quality of Life (QOL) models and patient specific parameters are combined to create a script for treatment planning. Treatment plans are optimised partly automatically (with QOL-weighted NTCP-based optimisation functions) and partly interactively (with conventional dose-based optimisation functions). The QOL score of the final plan can be calculated by combining the NTCP and QOL models, patient specific parameters and dose-volume histogram (DVH) data (see supplementary data for a detailed example).

QOL-weighted NTCP-guided plans were created in less time. On average, 5 optimisation sequences with manual adjustments (30 min) were needed to arrive at a balanced acceptable plan and an additional 10 optimisation sequences without further adjustments (60 min) were needed to maximise QOL. Resetting QOL-weighted NTCP-guided plans and then creating them again from scratch resulted in similar plans with similar QOL scores.

At least 98% of the PTVs received 95% of the prescribed dose in all plans and all patients and PTV D98 values were identical for the clinical plans and QOL-weighted NTCP-guided plans (Supplementary Table 5). With QOL-weighted NTCP-guided plans, the dose to the high risk PTVs was slightly less homogeneous and the dose to the intermediate risk PTVs was slightly less conformal (Supplementary Table 5). However, each plan complied with the primary planning objectives (Supplementary Table 2).

With QOL-weighted NTCP-guided plans, all OARs, except for the parotid glands, received a lower dose when compared with the clinical plans (Table 1). The largest reductions in dose were observed in OARs related to toxicities and symptoms that have a higher relative impact on QOL, such as the oral cavity and pharyngeal constructor muscles (related to dysphagia) and the glottic area and the arytenoids (related to speech problems). The relatively low impact on QOL of salivary function related toxicities and symptoms (Supplementary Table 3), caused the radiation dose to shift towards the parotid glands in the QOL-weighted NTCP-guided plans (Table 1). Due to this dose shift, NTCP values of toxicities and symptoms related to swallowing, speech, pain, and general complaints such as fatigue were all significantly lower with QOL-weighted NTCP-guided planning, at the expense of higher NTCP values for toxicities and symptoms related to salivary function (Table 2). The integral dose was significantly lower in QOL-weighted NTCP-guided plans because integral dose is related to fatigue, which is the symptom with the highest impact on QOL (Table 1; Supplementary Table 3). Still, absolute reductions in the NTCP of fatigue were limited as the options to reduce the integral dose were also limited (Table 2).

At 6 months, average predicted QOL was 72.8 and 73.5 points with the clinical plans and QOL-weighted NTCP-guided plans, respectively (Table 3). The imaginary zero dose plan was predicted to result in an average QOL score of 81.0. The average increase of 0.7 ± 0.5 (range: 0–1.8) QOL score with QOL-weighted NTCP-guided plans compared to the clinical plans constituted a gain of 8.6 ± 5.8 percent (range: 1.1–24.3 percent), given the highest QOL possible with imaginary zero dose plans (Fig. 2).

At later time points, the predicted increase of QOL with QOL-weighted NTCP-guided plans was greater: 0.9, 1.0 and 1.1 points at 12, 18 and 24 months, respectively (Table 3, Supplementary Table 6). Over time, it was estimated that 8.6 to 10.6 percent of radiotherapy-induced loss of QOL is potentially avoided by using QOL-weighted NTCP-guided planning instead of conventional VMAT planning (Table 3).

Discussion

To our knowledge, this was the first study to report on semi-automatic QOL-weighted NTCP-guided VMAT optimisation. We replaced the conventional EUD-based OAR optimisation functions in the treatment planning system by 80 NTCP-based objective functions corresponding to a wide range of toxicities and symptoms. Each objective function was weighted by its relative impact on QOL. This optimisation method demonstrated to be feasible and resulted in clinically acceptable treatment plans. Furthermore, we were able to evaluate and compare plans based on their expected QOL score. This allowed us to select the best plan in terms of expected QOL for individual patients. We demonstrated that QOL-weighted NTCP-guided plans resulted in a higher QOL compared to conventional EUD-based OAR optimisation. QOL was predicted to be better because during treatment optimisation, OAR related to the high-impact toxicities and symptoms were spared with priority, at the expense of increased dose in OAR related to lower-impact toxicities and symptoms.

Conventional EUD-based OAR optimisation is a labour-intensive procedure. After an initial trial optimisation run, the treatment planner gains a feeling of what the best starting point would be for the various OAR objectives and weights. Subsequently, the iterative process begins in search for optimal values and weights. This process is often very time consuming, subjective, and is expected to cause significant variability among treatment plans. Compared to the clinical plans, QOL-weighted NTCP-guided plans were produced in less time and were less dependent on the dosimetrists experience and yielded more consistent results as no trial-and-error adjustments were made to the OAR objectives and weights during optimisation. In the current implementation of this method a balance only needs to be found by the treatment planner between target coverage, general planning objectives and the risk of side effects. Therefore, this method could be regarded as a semi-automatic treatment planning approach.

Automated planning is nowadays a high-profile topic. Many different methods have been discussed in literature [11–14]. In most cases, a library of 'good' and 'representative' plans is available for learning the treatment planning system such that dose distributions can be predicted for new patients. In general, the drawbacks of these methods are: (1) new plans depend greatly on the quality of the plans available in the library; (2) the predicted plans are in most cases 'good' plans, but not necessarily optimal for individual patients; (3) no patient baseline and population outcome data is directly used to find the optimal solution for individual patients; (4) as time goes by, libraries are replenished with newer plans that incorporate no new knowledge as the newer plans themselves were also based on older knowledge; (5) although skill and experience of dosimetrists and radiation oncologists was used to create

Hans Paul van der Laan, A. van der Schaaf, L. Van den Bosch et al.

Radiotherapy and Oncology 162 (2021) 85–90

Table 3
Predicted Quality of Life.

	ZERO RT dose mean ± st. error	Conventional mean ± st. error	QOLO mean ± st. error	QOLO— Conventional mean ± st. error	p-value	QOLO gain/maximum achievable gain mean % ± st. deviation (range)
Predicted Quality of Life						
6 months	81.0 ± 1.4	72.8 ± 1.7	73.5 ± 1.7	0.7 ± 0.1	<0.001	8.6 ± 5.8 (range: 1.1–25.5)
12 months	80.1 ± 1.4	70.8 ± 1.9	71.7 ± 1.8	0.9 ± 0.1	<0.001	9.7 ± 6.5 (range: 1.7–29.5)
18 months	80.3 ± 1.2	70.2 ± 1.6	71.2 ± 1.6	1.0 ± 0.1	<0.001	10.4 ± 6.6 (range: 2.6–30.1)
24 months	79.8 ± 1.3	69.1 ± 1.7	70.2 ± 1.7	1.1 ± 0.1	<0.001	10.6 ± 6.6 (range: 3.1–30.6)

Abbreviations: Conventional = Clinical plans optimised by using equivalent uniform dose OAR based objectives and weights; QOLO = Quality of Life-optimised (QOL-weighted NTCP-guided); QOLO gain/maximum achievable gain = (QOLO−Conventional)/(ZERO RT dose−Conventional); ZERO RT dose = QOL prediction assuming a dose of zero Gy in all OAR.
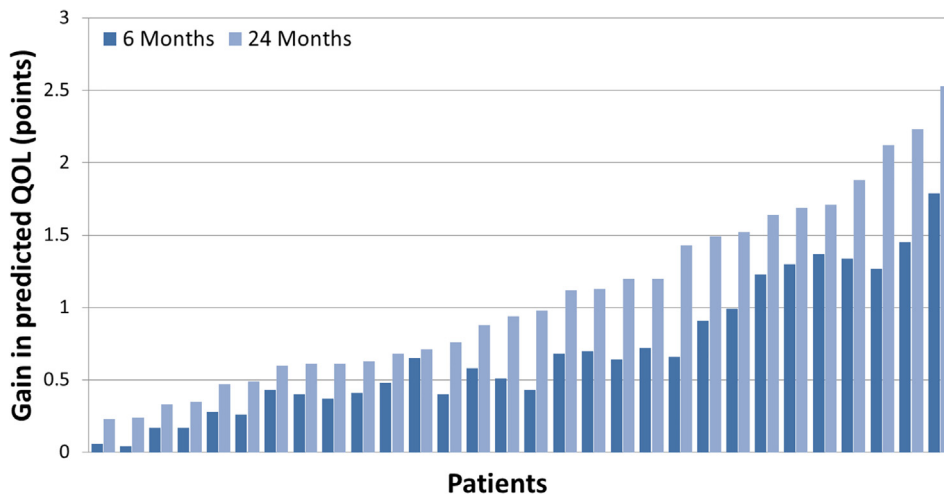

Fig. 2. Histograms represent differences in predicted Quality of Life (QOL) at 6 and 24 months for individual patients. The bars show the gain in QOL points with QOL-weighted NTCP-guided planning relative to conventional VMAT planning.

the initial library plans, as time goes by and automatic planning becomes the norm, these skills and experience will disappear and it becomes harder over time to incorporate new knowledge and improve the system; (6) in many cases, automatically created plans need some additional fine tuning (automated planning is seldom fully-automated planning) and it therefore relies on skilled and experienced treatment planners. The currently proposed semi-automatic QOL-weighted NTCP-guided planning method does not rely on a library of previous plans but instead is directly based on population and treatment outcome data. This data can be updated regularly as part of a rapid learning health care system and can be incorporated in the NTCP models used for continuous improvement of treatment optimisation.

With any dose optimisation method and with QOL-weighted NTCP-guided planning, dose can be redistributed to OARs related to toxicities or symptoms that are currently not accounted for. More-over, the analysis of quantitative plan quality measures (Supplementary Table 5) showed that, with QOL-weighted NTCP-guided planning, the dose to the high risk PTV was less uniform and that the dose was less conformal to the intermediate risk PTV. The impact of these redistributions of dose is not yet clear. Although the current NTCP profile includes 20 toxicities and symptoms [3,7], a number of toxicities, such as hearing problems, osteoradionecrosis, tube-feeding dependence, cerebrovascular accidents and others, have not yet been included. Neither have objective measures such as sali-vary flow measurements or assessments of swallowing function or

aspiration by means of video-fluoroscopy. The NTCP profile is contin-uously expanding and additional NTCP models will be added as new optimisation functions in the future.

QOL-weighted NTCP-guided planning has the potential to improve treatment for individual patients much more than con-ventional EUD-based OAR optimisation because multivariable NTCP-models allow for combining and balancing multiple factors into a single objective. Moreover, this method offers the opportu-nity to consider concomitant treatment and patient individual baseline characteristics such as existing toxicities and symptoms which further enables personalised treatment for patients.

For the current study we used a development version of the RayStation treatment planning system, as the method we used is not yet available in the clinical version. As this is also the case for treatment planning systems of other vendors, widespread implementation of QOL-weighted NTCP-guided planning is antici-pated, but now is not yet possible.

When applying the QOL model for treatment plan evaluation, it should be noted that many factors contribute to the QOL of individ-ual patients. Therefore, the QOL predictions should not be regarded as absolute QOL predictions for individual patients. Instead, the QOL predictions are an effective means to compare different treat-ment plans, as all other patient specific circumstances are constant when alternative treatment plans are considered.

Use of the current method is limited to patients for which the NTCP and QOL models were created, i.e., patients receiving

definitive (chemo)radiotherapy for HNC. Patients receiving postoperative radiotherapy sustain specific toxicities and symptoms that have not yet been considered. These will be added in future projects and are expected to increase QOL for these patient groups as well.

After clinical implementation we aim to perform a clinical validation study to assess whether QOL-weighted NTCP-guided planning does indeed increase the QOL of patients.

In conclusion, we demonstrated that QOL-weighted NTCP-guided planning is feasible as a semi-automatic treatment planning method. Plans were personalised, produced in less time, less dependent on the treatment planner experience and yielded more consistent results. It prioritises the sparing of OAR related to toxicities and symptoms that have the highest impact on QOL, considers patient specific baseline characteristics and can provide a systematic increase in predicted QOL.

Conflicts of interest statement

The department of Radiation Oncology of the University Medical Center Groningen has research collaborations with RaySearch.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2021.06.035.

References

[1] Wu Q, Manning M, Schmidt-Ullrich R, Mohan R. The potential for sparing of parotids and escalation of biologically effective dose with intensity-modulated radiation treatments of head and neck cancers: a treatment design study. Int J Radiat Oncol Biol Phys 2000;46:195–205. https://doi.org/10.1016/S0360-3016(99)00304-1.

[2] van der Laan HP, Gawryszuk A, Christianen MEMC, Steenbakkers RJHM, Korevaar EW, Chouvalova O, et al. Swallowing-sparing intensity-modulated radiotherapy for head and neck cancer patients: treatment planning optimization and clinical introduction. Radiother Oncol 2013;107:282–7. https://doi.org/10.1016/j.radonc.2013.05.004.

[3] Van den Bosch L, van der Schaaf A, van der Laan HP, Hoebers FJP, Wijers OB, van den Hoek JGM, et al. Comprehensive toxicity risk profiling in radiation therapy for head and neck cancer: a new concept for individually optimised treatment. Radiother Oncol 2021;157:147–54. https://doi.org/10.1016/j.radonc.2021.01.024.

[4] Sharabiani M, Clementel E, Andratschke N, Hurkmans C. Generalizability assessment of head and neck cancer NTCP models based on the TRIPOD criteria. Radiother Oncol 2020;146:143–50. https://doi.org/10.1016/j.radonc.2020.02.013.

[5] Kierkels RGJ, Korevaar EW, Steenbakkers RJHM, Janssen T, van't Veld AA, Langendijk JA, et al. Direct use of multivariable normal tissue complication probability models in treatment plan optimisation for individualised head and neck cancer radiotherapy produces clinically acceptable treatment plans. Radiother Oncol 2014;112:430–6. https://doi.org/10.1016/j.radonc.2014.08.020.

[6] Dahele M, Tol JP, Vergeer MR, Jansen F, Lissenberg-Witte BI, Leemans CR, et al. Is the introduction of more advanced radiotherapy techniques for locally-advanced head and neck cancer associated with improved quality of life and reduced symptom burden? Radiother Oncol 2020;151:298–303. https://doi.org/10.1016/j.radonc.2020.08.026.

[7] van der Laan HP, Van den Bosch L, Schuit E, Steenbakkers RJHM, van der Schaaf A, Langendijk JA. Impact of radiation-induced toxicities on quality of life of patients treated for head and neck cancer. Radiother Oncol 2021;160:47–53. https://doi.org/10.1016/j.radonc.2021.04.011.

[8] Grégoire V, Ang K, Budach W, Grau C, Hamoir M, Langendijk JA, et al. Delineation of the neck node levels for head and neck tumors: A 2013 update. DAHANCA, EORTC, HKNPCSG, NCIC CTG, RTOG, TROG consensus guidelines. Radiother Oncol 2014;110:172–81. https://doi.org/10.1016/j.radonc.2013.10.010.

[9] 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;117:83–90. https://doi.org/10.1016/j.radonc.2015.07.041.

[10] Qi XS, Semenenko VA, Li XA. Improved critical structure sparing with biologically based IMRT optimization. Med Phys 2009;36:1790–9. https://doi.org/10.1118/1.3116775.

[11] McIntosh C, Welch M, McNiven A, Jaffray DA, Purdie TG. Fully automated treatment planning for head and neck radiotherapy using a voxel-based dose prediction and dose mimicking method. Phys Med Biol 2017;62:5926–44. https://doi.org/10.1088/1361-6560/aa71f8.

[12] Hussein M, Heijmen BJM, Verellen D, Nisbet A. Automation in intensity modulated radiotherapy treatment planning-a review of recent innovations. Br J Radiol 2018;91. https://doi.org/10.1259/bjr.20180270.

[13] Moore KL. Automated radiotherapy treatment planning. Semin Radiat Oncol 2019;29:209–18. https://doi.org/10.1016/j.semradonc.2019.02.003.

[14] Bijman R, Sharfo AW, Rossi L, Breedveld S, Heijmen B. Pre-clinical validation of a novel system for fully-automated treatment planning. Radiother Oncol 2021;158:253–61. https://doi.org/10.1016/j.radonc.2021.03.003.