Differences between planned and delivered dose for head and neck cancer, and their consequences for normal tissue complication probability and treatment adaptation

Heukelom, J.; Kantor, M.E.; Mohamed, A.S.R.; Elhalawani, H.; Kocak-Uzel, E.; Lin, T.; ... ; Sonke, J.J.
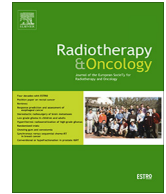
Citation

Original Article

Differences between planned and delivered dose for head and neck cancer, and their consequences for normal tissue complication probability and treatment adaptation

Jolien Heukelom [a], Michael E. Kantor [b], Abdallah S.R. Mohamed [c], Hesham Elhalawani [c], Esengul Kocak-Uzel [c,d], Timothy Lin [c], Jinzhong Yang [b], Michalis Aristophanous [b,e], Coen R. Rasch [f], Clifton David Fuller [c], Jan-Jakob Sonke [a,*]

[a] Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; [b] Department of Radiation Physics; [c] Division of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, USA; [d] Department of Radiation Oncology, SBU Sisli Etfal Teaching and Research Hospital, İstanbul, Turkey; [e] Department of Radiation and Cellular Oncology, The University of Chicago, Chicago, USA; and [f] Department of Radiation Oncology, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

ARTICLE INFO

ABSTRACT

Background and purpose: Anatomical changes induce differences between planned and delivered dose. Adaptive radiotherapy (ART) may reduce these differences but the optimal implementation is insufficiently clear. The aims of this study were to quantify the difference between planned and delivered dose in HNC patients, assess the consequential difference in normal tissue complication probability (ΔNTCP) and to explore the value of ΔNTCP as an objective selection strategy for ART.

Materials and methods: For 52 patients, daily doses were accumulated to estimate the delivered dose. The difference from planned dose was analyzed for CTVs and 9 organs-at-risk (OAR). ΔNTCP was calculated for xerostomia, dysphagia, parotid gland dysfunction and tube feeding dependency at 6 months. ART was deemed necessary if ΔNTCP was >5%. The positive predicted value (PPV) was calculated for identification of ART-patients by clinical judgement, and ΔNTCP at fraction 10 and 15.

Results: ΔNTCP >5% was seen five times for dysphagia and twice for the other toxicities. Only 5/9 patients with any ΔNTCP >5% clinically received ART, although ART had been done for 13/52 patients (PPV: 0.38). PPV was 0.86 and 0.75 for accumulated dose at fraction 10 and 15, respectively, using a ΔNTCP cut-off for the allocation of ART of 5%. Using other ΔNTCP cut-offs did not substantially improve PPV. With this cut-off the negative predictive value was 0.93 for ΔNTCP method of fraction 10 and fraction 15, and 0.90 for clinical judgement.

Conclusion: To identify patients accurately for ART, NTCP calculations based on the dose differences between planned and delivered dose at fraction 10 are superior to clinical judgement.

© 2019 Published by Elsevier B.V. Radiotherapy and Oncology 142 (2020) 100–106

Head and neck image-guided radiotherapy studies have demonstrated considerable geometric uncertainties, such as setup errors, posture changes, weight-loss and tumor shrinkage during fractionated radiotherapy [1,2]. These geometrical changes impact the dose distribution such that the actual delivered dose is different from the planned dose [3,4], impacting organs-at-risk (OAR) dose(s) to clinically meaningful degrees.

In adaptive radiotherapy (ART), one or more new radiotherapy plans are designed during the course of treatment in order to account for anatomical changes and thereby improve target

coverage, spare OAR and minimize toxicity. To date, at some facilities, patient selection for ART initiation is predicated on an individual decision by the treating physician to replan the patient, using expert opinion after visual inspection of one or more repeat images. Consequently, the physician determination is often based on heuristic, rather than quantitative indicators of improved dose distribution, resulting in an ART decision subject to inter-observer variability. Alternatively, many centers replan all patients at a fixed time-point, typically mid-therapy, to account for the larger part of anatomical changes, which are known to occur during the first half of treatment [5,6]. However, prospective data suggest that, while a majority of patients on trial did not have significant improvement from being replanned on a fixed schedule, a subset had substantive and clinically meaningful alteration [7]. Selection

based on differences between delivered dose and planned dose to targets and OARs afford systematic "triggers" of plan adaptation, and are more objective and robust (or, at least, quantifiable, and thus iteratively modifiable), precluding observer-dependent "guesses" as to the need for replanning and preventing inefficient resource overutilization embodied in "one-size-fits-all" strategies. Potentially the incorporation of the consequential difference in normal tissue complication probability (ΔNTCP) represents an improved method to select patients for ART, but objective data are lacking. Therefore, the specific aims of this study were to:

1) quantify the difference between the planned and delivered radiotherapy dose in head and neck cancer patients based on serial CT-on-rails imaging.
2) characterize organ-at-risk (OAR) delivered and planned radiotherapy dose discrepancies to assess potential consequential differences in NTCP.
3) determine positive predictive value (PPV) for ART based upon clinical judgement during treatment vs. a novel metric, namely ΔNTCP$_{F10-F15}$.

Materials and methods

Retrospective data was collected from consecutive patients treated for head and neck cancer with (chemo-) radiotherapy at MD Anderson Cancer Center, treated 2007–2013 with daily CT-on-Rails IGRT (CToR) (ExaCT, GE Healthcare, Waukesha, WI) [1]. All CToR scans were made for position verification, and anonymized before use in this study, therefore, no additional patient consent was required. The treatment planning CT was made after immobilizing patient with thermoplastic head and shoulder mask. Slice thickness ranged 2.5–3.75 mm. An expert attending head and neck radiation oncologist contoured all target lesions and OAR on the planning CT according to institutional head-and-neck contour guidelines. The margin from clinical to planning target volume was 3–5 mm. Pinnacle treatment planning system, (version 9.1 Philips Medical Systems, Eindhoven, The Netherlands) was used for delineation and treatment plan design. All plans were 9 beam step-and-shoot intensity modulated radiation therapy plans in adherence to the RTOG treatment guidelines. They were peer-reviewed [8–10]. For daily position verification, the patient was aligned to treatment position using skin marks on the linear accelerator, then the table turned 180 degrees and CToR scan was made in treatment position. The plan isocenter was related to the daily CT imaging using skin marks. This process was estimated to introduce a maximum uncertainty of 1 mm in the plan isocenter localization on daily imaging. The daily CT was automatically registered to the planning CT using in-house developed software, leading to a proposed alignment of the two scans, and a corresponding set-up correction [11]. The radiation therapists reviewed the alignment and made manual adjustments as they deemed necessary and per the instructions of the radiation oncologist.

Planned dose was the dose per region of interest (ROI) as derived from the planning CT with the original plan. To calculate the delivered dose, the initial treatment plan was recalculated for each day using the daily CT scan. If the patient position had been adjusted, the treatment isocenter was moved accordingly before recalculation of dose. The new dose distribution file was mapped back to the planning CT using a deformable deformation vector field (DVF), created with software validated for this purpose (Admire 1.04, Elekta AB, Stockholm, SE) [12]. In the planning CT, the original contours were used to calculate the daily dose to that organ. If there were missing daily CTs, the CTs from the closest two fractions were used for linear interpolation to generate the missing fractionated dose. If first or last fractions were missing, nearest neighbor extrapolation was used.

D1 and D99 were calculated for the primary, high risk and elective clinical target volume (CTV), the affix '_sub' identified volumes from which the higher dose levels had been subtracted. To complement the dataset, additional OAR contours were created using an atlas based strategy (Admire (Elekta AB, Stockholm, Sweden)) in which the atlases were based on institutional head-and-neck contour guidelines, and all the generated contours were visually inspected and corrected by a radiation oncologist [EK] [13].

NTCP models

The absolute dose differences between planned and delivered dose were calculated for all targets and OAR based on the original treatment plan. Additionally, NTCP models were used to assess the difference in expected toxicity for the following end points: patient reported moderate to severe xerostomia at 6 months, Xer6m [14], physician rated feeding tube dependency at 6 months, Tube6m [15], physician rated decreased salivary flow using the mean dose model [16] and dysphagia i.e., grade 2–4 swallowing dysfunction according to the RTOG/EORTC Late Radiation Morbidity Scoring Criteria, physician rated 6 months after treatment (Appendix A) [17].

The models were chosen based on provisory national consensus from the Dutch head and neck radiation oncology committee, and are NTCP models currently implemented for therapeutic selection of patients for proton-therapy [18]. Appendix B states how was dealt with missing input parameters for NTCP models.

Statistical analyses

Differences in either parameters of the accumulated dose or NTCP based on delivered dose were tested for statistical significance for all patients, using Wilcoxon signed rank test, corresponding p-values are for 2-tailed significance, unless specified otherwise. A Bonferroni correction was done regarding the number of volumes/organs at risk that were tested; a p-value $< (0.05/10)$ was considered statistically significant for the OAR, for the CTVs this was a p-value $< (0.05/5)$. Calculations were done using IBM SPSS statistics v22 or JMP v13 (SAS Institute, Cary, NC, USA). To predict ΔNTCP at the end of treatment using the dose difference until fraction 10 (F10) and 15 (F15), the accumulated dose on that day was scaled to the full dose prior to the NTCP calculation. Currently, the Dutch Society for Radiation Oncology has set the ΔNTCP threshold for allocation to proton treatment at 10% and 5% for grade II and III complications respectively. ART is less costly than proton treatment. It was predetermined at a consensus investigator meeting that patients should be treated with ART more readily than with proton therapy. Therefore, a 5% threshold for ΔNTCP at the end of treatment was determined to be clinically relevant in this study. The positive predictive value, negative predictive value, sensitivity and specificity for identifying these patients were calculated for clinical judgement and ΔNTCP based on dose differences at F10 or F15. Different thresholds for ΔNTCP at those time points were assessed for their predictive accuracy). For the purposes of this study, a true-positive represented a scenario in which replanning was performed and >5% ΔNTCP was observed at end-treatment; a false-positive indicated a "wasted replan" wherein ART was performed, but no ΔNTCP >5% was observed; a true-negative indicated a non-replanned case where ΔNTCP <5%; and false-negative denoted an unaltered plan with end-therapy ΔNTCP >5%. The positive predictive value was considered the most important for patient allocation to ART.

In order to determine the relative efficacy of a specific replanning regimen receiver operator characteristic curve (ROC) analysis was performed. A method was deemed better than another if the asymptotic significance of the difference between the area-

under-the-curve (AUC) was statistically different by the Mann-Whitney U test. For chance an AUC of 0.5 was used.

Results

Eighty-seven patients with daily intra-treatment image guided CToR were identified, of whom fifty-two were eligible for inclusion. The most common reason for exclusion was missing isocenter shift information. Demographics and dose prescriptions were tabulated (Supplementary Table 1 and 2 respectively). 13 patients clinically received ART. There were no standardized criteria for when ART was performed. Typically, the treating physician would adapt the treatment plan if any of the following occurred: weight loss >10%, tumor shrinkage, contour alterations within the treatment mask or deteriorated mask fitting. The earliest ART was after 8 fractions, the last was after 23 fractions (Supplementary Table 2).

For the whole cohort, the D99 and D95 were statistically different for all measured CTVs, with a wide range of observed differences; nonetheless, median $CTV_{planned}$–$CTV_{delivered}$ dose difference did not exceed ±5% of prescription dose for any patient (Fig. 1, Supplementary Table 3). Maximum dose to the CTVs was not statistically significantly different for delivered compared to planned dose. On an individual basis, large differences (>15%) between planned and delivered dose were occasionally seen: for example there were four patients for whom the D99 of the delivered dose to the primary tumor was more than 15% less than planned (Fig. 1; Supplementary Table 2: patient 9, 12, 22 and 38).

The difference between planned and delivered dose was calculated for mean and maximum dose in nine OARs. Additional atlas-based contours were used for nine OAR (Table 1). There was a significant difference between mean planned and delivered dose to the larynx (p = 0.006) and mandible (p = 0.001), with small median differences: 0.84 Gy and −0.53 Gy respectively. The

delivered maximum dose was significantly different from the planned dose for brainstem ($p < 0.001$) and spinal cord (p = <0.001). Median ΔD1 for these OAR was 1.03 Gy and 0.96 Gy respectively. D1 was also significantly lower for planned than delivered dose to the mandible, contralateral and ipsilateral parotid gland, with median dose differences <0.5 Gy.

For all patients combined, there was no statistical significant difference between the NTCP for planned and delivered dose regarding dysphagia (p = 0.12), xer6m (p = 0.26) or parotid gland dysfunction (p = 0.88). For tube6m there was a statistical significant difference (p = 0.028), but with low clinical impact, as the median absolute ΔNTCP was 0.13%. (Fig. 2).

However, large individual ΔNTCP were observed. At end treatment, a clinically relevant ΔNTCP (>5%) was found 11 times for all models combined (five times for dysphagia, twice for each of the other toxicities), in nine patients (17%). ΔDose by OARs is summarized in Table 1.

Only 5/9 patients with any ΔNTCP >5% at the end of treatment clinically received ART, although ART had been done for 13/52 patients (PPV: 0.38). PPV was higher for our models: 0.86 and 0.75 at F10 and F15 respectively using a ΔNTCP threshold at that time of 5% (Table 2). For F15, using a 4% ΔNTCP slightly improved PPV, but it was best at F10 with the 5% ΔNTCP threshold (Table 2). True positive at F10 at 5% threshold not only identified the correct patients, but also the correct toxicity (data not shown). For 1 patient ΔNTCP at F10 was >5% in two models, of which only one had ΔNTCP at last fraction of >5%. At the 5% threshold the negative predictive value was 0.93 for F10 and F15, and 0.90 for clinical judgement. In other words, of the 13 patients that received ART based on clinical judgement, only 5 actually benefitted (true-positive), whilst 4 patients who had a ΔNTCP >+5% were missed (false-negative). If F10 would have been used to allocate patients to ART with a threshold of >5%ΔNTCP, six true positives would
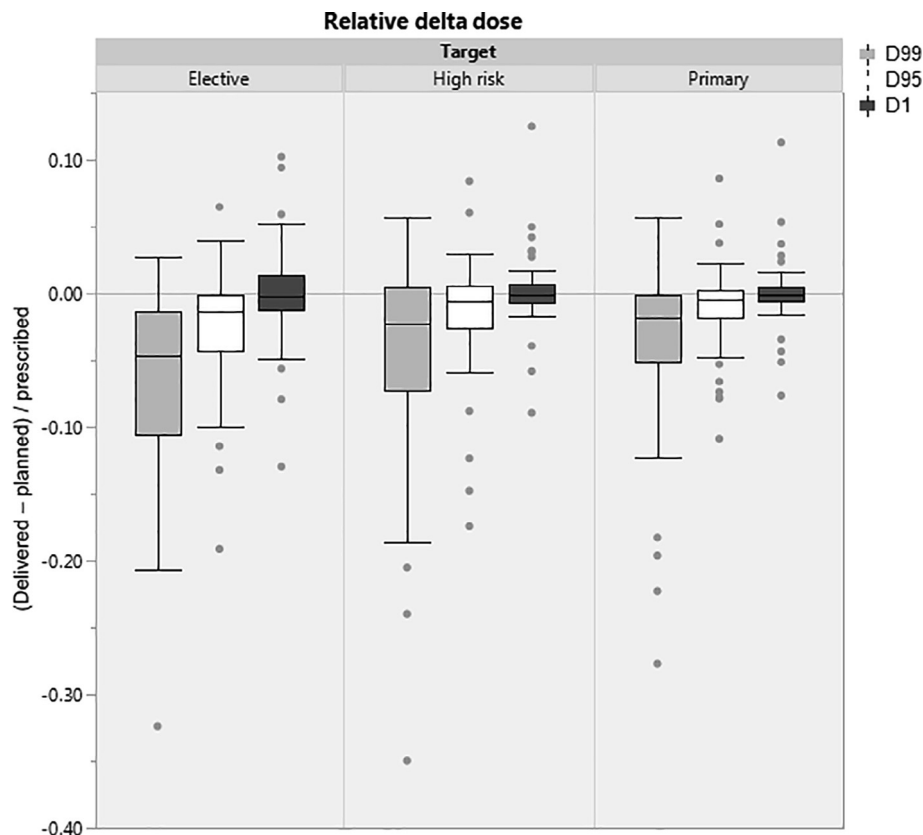

Fig. 1.

Table 1

Delta (delivered – planned) dose to OAR. The parotid with the highest mean planned dose was referred to as ipsilateral. The ipsilateral submandibular gland was named according to the parotid gland. Manually contoured OARs were supplemented by automatic contours when possible. It could not be done for chiasm. P^W p-values from Wilcoxon signed rank test for two dependent samples, 2-tailed significance. A Bonferroni corrected p-value <0.005 was considered statistically significant.

| Delta dose; (delivered – planned dose)		Delta mean dose (Gy)			Delta D1 (Gy)		
Structure	N (Manual)	Median	Range	P^W	Median	Range	P^W
Brainstem	52 (50)	0.25	−3.92 −3.78	0.14	1.03	−4.72 −0.59	<0.001
Chiasm	31 (31)	0.62	−12.67 −5.72	0.95	1.07	−8.20 −8.67	0.71
Esophagus	52 (7)	0.00	−4.90 −4.44	0.29	0.00	−8.25 −6.29	0.85
Larynx	50 (12)	0.84	−8.14 −10.93	0.006	0.10	−21.79 −9.94	0.36
Mandible	51 (18)	−0.53	−6.24 −2.93	0.001	−0.49	−8.26 −6.28	0.001
Parotid Gland Contralateral	52 (49)	−0.21	−5.42 −7.05	0.31	−0.47	−11.52 −4.21	0.04
Parotid Gland Ipsilateral	50 (43)	0.01	−6.90 −5.63	0.82	−0.44	−5.91 −4.09	0.01
Spinal_cord	51 (51)	0.71	−1.52 −4.08	<0.001	0.96	−3.20 −8.09	<0.001
Submandibular Gland Contralateral	52 (15)	−0.010	−5.38 −5.15	0.68	−0.14	−13.29 −8.05	0.15
Submandibular Gland Ipsilateral	52 (6)	−0.003	−4.78 −7.37	0.99	−0.25	−18.45 −7.72	0.14
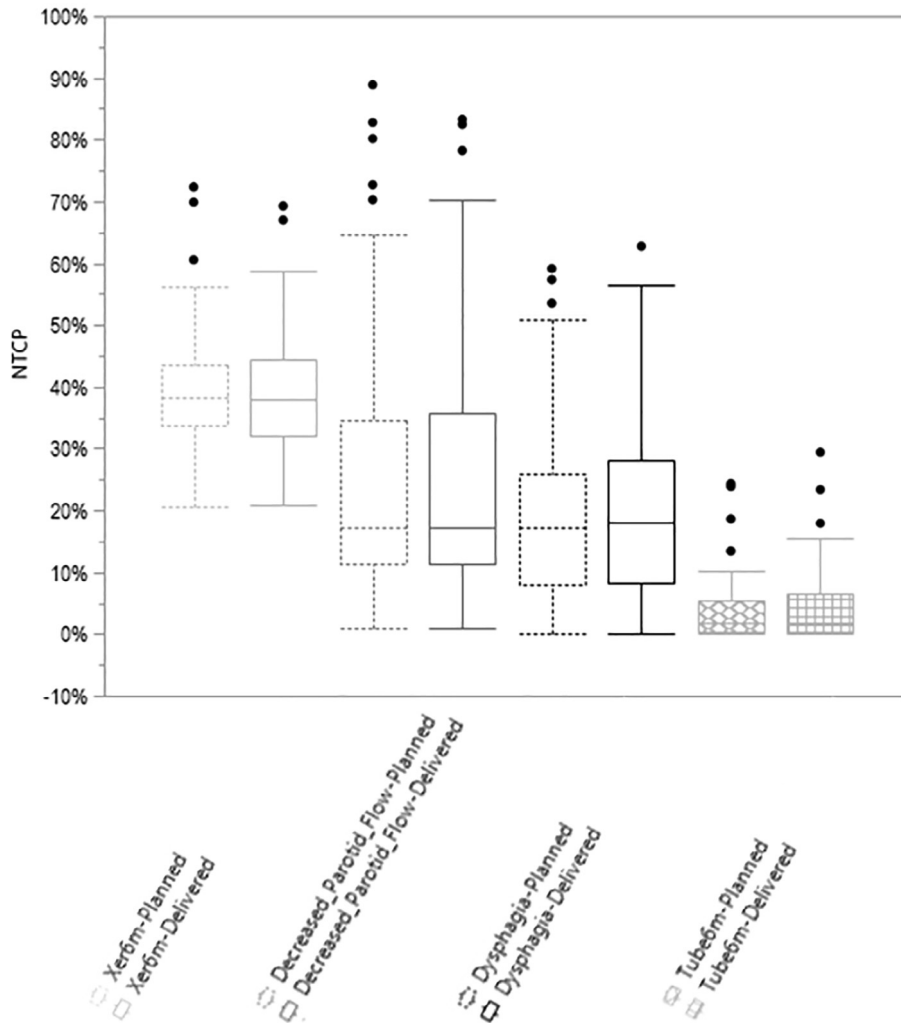
Fig. 2.

Table 2
Prediction of final NTCP using NTCP at F10 or F15 using various ΔNTCP thresholds. The percentages indicate the decision to adapt treatment for every patient that has a predicted ΔNTCP of x% or higher in any of the four NTCP models, based on the dose difference at Fx, scaled to a full-treatment length.

	NTCP whole treatment	Clinical	F10 1%	F10 2%	F10 3%	F10 4%	F10 5%	F10 6%	F10 7%	F10 8%	F10 9%	F10 10%	F10 11%	F10 12%
ART for # patients	9	13	26	19	15	9	7	5	5	4	4	4	3	2
True positive		5	8	8	8	7	6	4	4	3	3	3	2	2
False positive		8	18	11	7	2	1	1	1	1	1	1	1	0
True negative		35	25	32	36	41	42	42	42	42	42	42	42	43
False negative		4	1	1	1	2	3	5	5	6	6	6	7	7
Sensitivity		0.56	0.89	0.89	0.89	0.78	0.67	0.44	0.44	0.33	0.33	0.33	0.22	0.22
Specificity		0.81	0.58	0.74	0.84	0.95	0.98	0.98	0.98	0.98	0.98	0.98	0.98	1.00
PPV		0.38	0.31	0.42	0.53	0.78	0.86	0.80	0.80	0.75	0.75	0.75	0.67	1.00
NPV		0.90	0.96	0.97	0.97	0.95	0.93	0.89	0.89	0.88	0.88	0.88	0.86	0.86

	F15 1%	F15 2%	F15 3%	F15 4%	F15 5%	F15 6%	F15 7%	F15 8%	F15 9%	F15 10%	F15 11%	F15 12%
ART for # patients	29	21	12	10	8	7	6	4	4	4	2	1
True positive	9	9	8	8	6	5	4	3	3	3	2	1
False positive	20	12	4	2	2	2	2	1	1	1	0	0
True negative	23	31	39	41	41	41	41	42	42	42	43	43
False negative	0	0	1	1	3	4	5	6	6	6	7	8
Sensitivity	1.00	1.00	0.89	0.89	0.67	0.56	0.44	0.33	0.33	0.33	0.22	0.11
Specificity	0.53	0.72	0.91	0.95	0.95	0.95	0.95	0.98	0.98	0.98	1.00	1.00
PPV	0.31	0.43	0.67	0.80	0.75	0.71	0.67	0.75	0.75	0.75	1.00	1.00
NPV	1.00	1.00	0.98	0.98	0.93	0.91	0.89	0.88	0.88	0.88	0.86	0.84

Abbreviations: ART: adaptive radiotherapy; F: fraction; NPV: negative predictive value; NTCP: normal tissue complication probability; PPV: positive predictive power; #: number of.
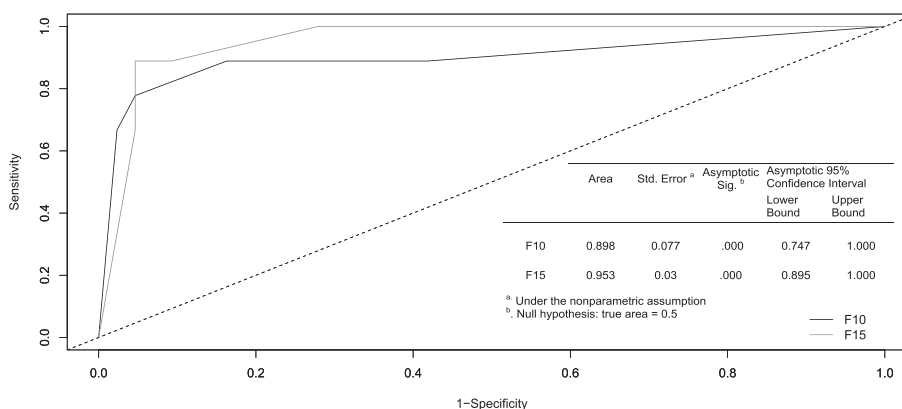

	Area	Std. Error [a]	Asymptotic Sig. [b]	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
F10	0.898	0.077	.000	0.747	1.000
F15	0.953	0.03	.000	0.895	1.000

[a] Under the nonparametric assumption
[b] Null hypothesis: true area = 0.5

Fig. 3.

have been identified, 'unnecessary' ART would have been done for one rather than eight patients, and only three false negatives would have occurred.

ROC analysis (Fig. 3) demonstrated substantively improved discrimination of cases by our model compared to chance. Both F10 and F15 had AUCs >0.85, they were not statistically different from each other.

Discussion

This study describes a uniquely large retrospective cohort of patients treated with image guided radiotherapy based on CT-on-Rails. The big advantage of CT-on-Rails is that dose reconstruction can be done directly and does not involve an additional computational step, which is required for recalculations based on cone-beam-CT. Furthermore, the image quality of CT-on-Rails is superior to that of CBCT. The drawback is that the sample size does not allow internal validation and that there is no global reference dataset available for external validation. The patient selection for treatment on this limited capacity device may pose a bias towards patients with expected anatomical changes and thus larger NTCP differences between calculations based on planned and delivered dose, but ART was done in 25% of the cases, which is not extraordinary. For all CTVs the D95 of delivered dose was significantly lower than the planned dose. However, the relative median difference in dose was always ≤2%. This is in agreement with a study by Graff et al. [4], who showed a difference in D95 with delivered dose generally less than 5% different from planned dose. In a group of 11 patients with ART, Duma et al. also showed the PTV was covered by the prescribed dose for all patients, even prior to ART [20].

For OAR, a significant difference was only seen between mean planned and delivered dose to the larynx ($p = 0.007$) and mandible ($p = 0.001$), although the median differences were small and presumably of no clinical consequence. Changes in dose to the mandible are more likely be due to rotational errors than to true anatomical changes [21,22]. For the other OAR, including the parotids, no difference in mean dose was found. These results are more optimistic than those in the review by Brouwer et al. [23], who found an average delta dose for the parotids of 2.2 Gy ± 2.7 Gy.

An alternative hypothesis is that the weekly images that are used in most studies transform random errors into systemic errors by recalculating them for several fractions, thereby overestimating their effect on the final delivered dose. To our knowledge, there are no studies available with daily CT to compare our results to. There are a few smaller studies which also show small differences between planned and delivered dose to the OAR, so our results are not unique [20]. Duma et al. found an average delta Dmean larynx of 1.3% higher than planned dose in a group of 10 patients, similar to our results [24]. With small differences in dose to the OAR, it is logical that at group level ΔNTCP is negligible. However, similar to the Schwartz study, individual changes can be substantial [7]. Moreover, our study showed NTCP modelling at F10 or F15 is significantly better than chance at predicting the need for ART. Therefore, NTCP calculations provide a valuable tool to aid physicians in the future to test the probable gain of ART based on the OAR and NTCP. The difference in predictive accuracy for F10 and F15 are similar in our study, but this may be due to limitations in sample size. Analogue to our clinical decisions, Muller et al. performed a dose summation and replanning study in which they conclude that for photon based treatment, OAR are more important to decide for ART than the target volumes, as coverage stays good over the course of treatment [25]. Although at group level, the same can be said for the current study, Fig. 1 also shows large incidental individual differences between planned and delivered dose to targets, which could very well justify ART. Meaningful use of ART improves quality of life and seems most effective early in treatment, as most anatomical changes occur then [5,6,22,24,25], which is especially true for centers where ART efforts are labor intensive and time consuming.

To our knowledge, this study has the largest cohort of HNC patients with daily CT images available to assess changes in anatomy over treatment, as opposed to the often-used conebeam CT (CBCT). CT has advantage over CBCT since it suffers less from scatter, hence producing a higher contrast-to-noise ratio and reliable Hounsfield units. Thus in our study, the consequential differences between planned and delivered dose could be assessed with high granularity. Moreover, only few earlier publications have reported on these dosimetric differences and NTCP [23].

Possible limitations of this study include the heterogeneity of the patient cohort, its retrospective nature and associated potential bias of the patient selection. The ΔNTCP incidences reported in this paper might therefore not be representative for other institutes. Because of insufficient power for subgroup analyzes, dose differences between planned and delivered dose were not analyzed for their correlation with pathology, primary site, tumor stage, concurrent treatment or HPV status. Although we did not have all parameters of every NTCP model available, the results were considered reliable, based on our calculations comparing different substitute values (Appendix). It is debatable whether NTCP models designed for planned dose are applicable for delivered dose, but they are the best available. A refit of NTCP models based on delivered dose rather than planned is not expected to substantially impact NTCP, as the absolute dose differences we found were small. One of the few studies addressing this subject was done by Hunter et al., who investigated planned and delivered dose in their ability to predict a decrease in measured salivary flow post treatment and found their predictive power was not significantly different [26]. In this work we used a scaling of the accumulated dose early in treatment to estimate the delivered dose at the end of treatment. Such an approach would underestimate the impact of anatomical changes if the anatomy progressively changes over the whole course of treatment as the changes that continue after the initial estimate are not accounted for. Usually, however, they are more pronounced in the initial stage of the treatment [5,6,26]. Using the simple scaling approach, our model already has a good positive predictive value of 0.86 at fraction 10. More refined extrapolation method of the initial accumulated dose to the end of treatment might improve the model.

In short, our study illustrates NTCP calculations are superior to clinical judgement in patient selection for ART. Further research should be directed at identification of patients who might benefit from ART based on normal tissue toxicity probability in addition to clinical judgement.

Funding

Declaration of Competing Interest

Dr. Heukelom reports grants from European Union 7th Framework, personal fees from Dutch Cancer Society, personal fees from Rene Vogels stichting, during the conduct of the study. Dr. Fuller reports grants from Elekta AB, during the conduct of the study; grants from National Institute for Dental and Craniofacial Research, grants from National Cancer Institute, outside the submitted work. Dr. Sonke reports grants from EU 7th framework program, during the conduct of the study; grants from Elekta AB, other from Elekta AB, outside the submitted work; In addition, Dr. Sonke has a patent Computed Tomography Scanning with royalties paid to Elekta AB, a patent Motion Artefact Reduction in CT Scanning issued to NA, a patent Methods and System for Protecting Critical Structures During Radiation Treatment with royalties paid to Elekta AB, and a patent Radiotherapy and Imaging Methods and Apparatus with royalties paid to Elekta AB.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2019.07.034.

References

[1] Barker JL, Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, et al. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol Biol Phys 2004;59:960–70. https://doi.org/10.1016/j.ijrobp.2003.12.024.

[2] Ho KF, Marchant T, Moore C, Webster G, Rowbottom C, Penington H, et al. Monitoring dosimetric impact of weight loss with kilovoltage (KV) cone beam CT (CBCT) during parotid-sparing IMRT and concurrent chemotherapy. Int J Radiat Oncol Biol Phys 2012;82:e375–82. https://doi.org/10.1016/j.ijrobp.2011.07.004.

[3] O'Daniel JC, Garden AS, Schwartz DL, Wang H, Ang KK, Ahamad A, et al. Parotid gland dose in intensity-modulated radiotherapy for head and neck cancer: is what you plan what you get?. Int J Radiat Oncol Biol Phys 2007;69:1290–6. https://doi.org/10.1016/j.ijrobp.2007.07.2345.

[4] Graff P, Kirby N, Weinberg V, Chen J, Yom SS, Lambert L, et al. The residual setup errors of different IGRT alignment procedures for head and neck IMRT and the resulting dosimetric impact. Int J Radiat Oncol Biol Phys 2013;86:170–6. https://doi.org/10.1016/j.ijrobp.2012.10.040.

[5] Bhide SA, Davies M, Burke K, McNair HA, Hansen V, Barbachano Y, et al. Weekly volume and dosimetric changes during chemoradiotherapy with intensity-modulated radiation therapy for head and neck cancer: a prospective

observational study. Int J Radiat Oncol Biol Phys 2010;76:1360–8. https://doi.org/10.1016/j.ijrobp.2009.04.005.

[6] Wang Z-H, Yan C, Zhang Z-Y, Zhang C-P, Hu H-S, Kirwan J, et al. Radiation-induced volume changes in parotid and submandibular glands in patients with head and neck cancer receiving postoperative radiotherapy: a longitudinal study. Laryngoscope 2009;119:1966–74. https://doi.org/10.1002/lary.20601.

[7] Schwartz DL, Garden AS, Shah SJ, Chronowski G, Sejpal S, Rosenthal DI, et al. Adaptive radiotherapy for head and neck cancer–dosimetric results from a prospective clinical trial. Radiother Oncol 2013;106:80–4. https://doi.org/10.1016/j.radonc.2012.10.010.

[8] Cardenas CE, Mohamed ASR, Tao R, Wong AJR, Awan MJ, Kuruvila S, et al. Prospective qualitative and quantitative analysis of real-time peer review quality assurance rounds incorporating direct physical examination for head and neck cancer radiation therapy. Int J Radiat Oncol Biol Phys 2017;98:532–40. https://doi.org/10.1016/j.ijrobp.2016.11.019.

[9] Rosenthal DI, Asper JA, Barker JL, Garden AS, Chao KSC, Morrison WH, et al. Importance of patient examination to clinical quality assurance in head and neck radiation oncology. Head Neck 2006;28:967–73. https://doi.org/10.1002/hed.20446.

[10] Garden AS, Dong L, Morrison WH, Stugis EM, Glisson BS, Frank SJ, et al. Patterns of disease recurrence following treatment of oropharyngeal cancer with intensity modulated radiation therapy. Radiat Oncol Biol 2013;85:941–7. https://doi.org/10.1016/j.ijrobp.2012.08.004.

[11] Court LE, Dong L. Automatic registration of the prostate for computed-tomography-guided radiotherapy. Med Phys 2003;30:2750–7. https://doi.org/10.1118/1.1608497.

[12] Qin A, Liang J, Han X, O'Connell N, Yan D. Technical Note: the impact of deformable image registration methods on dose warping. Med Phys 2018;45:1287–94. https://doi.org/10.1002/mp.12741.

[13] Morassut S. Imaging atlas of human anatomy—3rd edition. Clin Imaging 2004;28:314–5. https://doi.org/10.1016/j.clinimag.2004.04.022.

[14] Beetz I, Schilstra C, Van Der Schaaf A, Van Den Heuvel ER, Doornaert P, Van Luijk P, et al. NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors q. Radiother Oncol 2012;105:101–6. https://doi.org/10.1016/j.radonc.2012.03.004.

[15] Wopken K, Bijl HP, Van Der Schaaf A, Paul H, Der Laan V, Chouvalova O, 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.

[16] Dijkema TIM, Raaijmakers CP, Ten Haken RK, Roesink JM, Braam PM, Houweling AC, et al. Parotid gland function after radiotherapy: the combined Michigan and Utrecht experience. Int J Radiat Oncol Biol Phys 2010;78:449–53. https://doi.org/10.1016/j.ijrobp.2009.07.1708.

[17] 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.

[18] 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.

[20] Castadot P, Lee JA, Geets X, Grégoire V. Adaptive radiotherapy of head and neck cancer. Semin Radiat Oncol 2010;20:84–93. https://doi.org/10.1016/j.semradonc.2009.11.002.

[21] Müller BS, Duma MN, Kampfer S, Nill S, Oelfke U, Geinitz H, et al. Impact of interfractional changes in head and neck cancer patients on the delivered dose in intensity modulated radiotherapy with protons and photons. Phys Med 2015;31:266–72. https://doi.org/10.1016/j.ejmp.2015.02.007.

[22] Zhang P, Simon A, Rigaud B, Castelli J, Ospina Arango JD, Nassef M, et al. Optimal adaptive IMRT strategy to spare the parotid glands in oropharyngeal cancer. Radiother Oncol 2016;120:41–7. https://doi.org/10.1016/j.radonc.2016.05.028.

[23] Brouwer CL, Steenbakkers RJHM, Ph D, Langendijk JA, Ph D, Sijtsema NM, et al. Identifying patients who may benefit from adaptive radiotherapy : does the literature on anatomic and dosimetric changes in head and neck organs at risk during radiotherapy provide information to help ? n.d.; l.

[24] Yang H, Hu W, Wang W, Chen P, Ding W, Luo W. Replanning during intensity modulated radiation therapy improved quality of life in patients with nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys 2013;85:e47–54. https://doi.org/10.1016/j.ijrobp.2012.09.033.

[25] Zhao L, Wan Q, Zhou Y, Deng X, Xie C, Wu S. The role of replanning in fractionated intensity modulated radiotherapy for nasopharyngeal carcinoma. Radiother Oncol 2011;98:23–7. https://doi.org/10.1016/j.radonc.2010.10.009.

[26] Hunter KU, Fernandes LL, Vineberg KA, Mcshan D, Antonuk AE, Cornwall C, et al. Parotid glands dose e effect relationships based on their actually delivered doses: implications for adaptive replanning in radiation therapy of head-and-neck cancer. Radiat Oncol Biol 2013;87:676–82. https://doi.org/10.1016/j.ijrobp.2013.07.040.