Advances in Radiation Oncology (2020) **5**, 111-119

advances
in radiation oncology

www.advancesradonc.org

Scientific Article

Radiation-Induced Hypothyroidism After Radical Intensity Modulated Radiation Therapy for Oropharyngeal Carcinoma

Mona Kamal, MD, PhD,[a,d] Christopher Ryan Peeler, PhD,[b] Pablo Yepes, PhD,[b,e] Abdallah S.R. Mohamed, MD, MSc,[a,f,g] Pierre Blanchard, MD,[a] Steven Frank, MD,[a] Lei Chen, PhD,[a] Amit Jethanandani, MPH,[a] Rohit Kuruvilla, MD,[a] Benjamin Greiner, MD,[a] Jared Harp, MD,[a] Robin Granberry, BA,[a] Vivek Mehta, MD,[a] Crosby Rock, MD,[a] Katherine Hutcheson, PhD,[c] Carlos Cardenas, PhD,[b] G.Brandon Gunn, MD,[a] Clifton Fuller, MD, PhD,[a,g] and Dragan Mirkovic, PhD[b,*]

Departments of [a]Radiation Oncology, [b]Radiation Physics, and [c]Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, Texas; [d]Department of Clinical Oncology and Nuclear Medicine, Faculty of Medicine, Ain Shams University, Cairo, Egypt; [e]Department of Physics and Astronomy, Rice University, Houston, Texas; [f]Department of Clinical Oncology and Nuclear Medicine, Faculty of Medicine, University of Alexandria, Alexandria, Egypt; and [g]MD Anderson Cancer Center/UTHealth Graduate School of Biomedical Sciences, Houston, Texas

Received 15 January 2019; revised 15 August 2019; accepted 20 August 2019

Abstract

Purpose: To evaluate 2 published normal tissue complication probability models for radiation-induced hypothyroidism (RHT) on a large cohort of oropharyngeal carcinoma (OPC) patients who were treated with intensity-modulated radiation therapy (IMRT).

Methods and Materials: OPC patients treated with retrievable IMRT Digital Imaging and Communications in Medicine (DICOMs) data and available baseline and follow-up thyroid function tests were included. Mean dose (Dmean) to the thyroid gland (TG) and its volume

Sources of support: This work is directly supported by the Andrew Sabin Family Foundation; Dr Fuller is a Sabin Family Foundation Fellow. Family of Paul W. Beach is providing direct salary support for Dr Kamal. Amit Jethanandani is supported by the Dunagan MD Medical Education Fund through The University of Tennessee Health Science Center, College of Medicine. Drs Yepes, Peeler, and Mirkovic are supported by Cancer Research Institute of Texas (CPRIT) award RP160232.

Disclosures: Drs. Mohamed, Hutcheson, and Fuller received funding support from the National Institutes of Health (NIH)/National Institute for Dental and Craniofacial Research (1R01DE025248-01/R56DE025248-01) and NIH/ National Cancer Institute (NCI) Early Phase Clinical Trials in Imaging and Image-Guided Interventions Program (1R01CA218148-01). Dr Fuller receives federal grant and salary support from the NIH/NCI Head and Neck Specialized Programs of Research Excellence (SPORE) Developmental Research Program Award (P50CA097007-10) and a National Science Foundation (NSF), Division of Mathematical Sciences, Joint NIH/NSF Initiative on Quantitative Approaches to Biomedical Big Data (QuBBD) Grant (NSF 1557679); the NCI QUBBD (1R01CA225190-01); the NIH Big Data to Knowledge (BD2K) Program of the NCI Early Stage Development of Technologies in Biomedical Computing, Informatics, and Big Data Science Award (1R01CA214825-01); NCI Early Phase Clinical Trials in Imaging and Image Guided Interventions Program (1R01CA218148-01); a General Electric Healthcare/MD Anderson Center for Advanced Biomedical Imaging In-Kind Award; an Elekta AB/MD Anderson Department of Radiation Oncology Seed Grant; the Center for Radiation Oncology Research (CROR) at MD Anderson Cancer Center Seed Grant; and the MD Anderson Institutional Research Grant (IRG) Program and the Cancer center Support Grant Radiation Oncology/Cancer Imaging Program Seed Grant (5P30CA016672). Dr Fuller received industry grant support and speaker travel funding from Elekta AB.

* Corresponding author: Dragan Mirkovic, PhD; E-mail: dmirkovi@mdanderson.org

https://doi.org/10.1016/j.adro.2019.08.006

were calculated. The study outcome was clinical HT at least 6 months after radiation therapy, which was defined as grade ≥2 HT per Common Terminology Criteria for Adverse Events grading system (symptomatic hypothyroidism that required thyroid replacement therapy). Regression analyses and Wilcoxon rank-sum test were used. Receiver operating characteristic curves and area under the curve for the fitted model were calculated.

Results: In the study, 360 OPC patients were included. The median age was 58 years. Most tumors (51%) originated from the base of tongue. IMRT-split field was used in 95%, and median radiation therapy dose was 69.96 Gy. In the study, 233 patients (65%) developed clinical RHT that required thyroid replacement therapy. On multivariate analysis higher Dmean and smaller TG volume maintained the statistically significant association with the risk of clinical RHT ($P < .0001$). Dmean was significantly higher in patients with clinical RHT versus those without (50 vs 42 Gy, $P < .0001$). Patients with RHT had smaller TG volume compared with those without (11.8 compared with 12.8 mL, $P < .0001$). AUC of 0.72 and 0.66 were identified for fitted model versus for the applied Boomsma et al and Cella et al models, respectively.

Conclusions: Volume and Dmean of the TG are important predictors of clinical RHT and shall be integrated into normal tissue complication probability models for RHT. Dmean and thyroid volume should be considered during the IMRT plan optimization in OPC patients.

Introduction

Hypothyroidism is one of the radiation therapy (RT)-attributable side effects after curative treatment of head and neck cancers (HNC).[1-3] It has been reported that 19% to 53%[4,5] of the HNC patients develop hypothyroidism after RT, which negatively affects quality of life in HNC survivors.[6] This is of particular importance in the era of human papillomavirus (HPV) positive oropharyngeal carcinoma (OPC) with rapidly growing numbers of (young) survivors who live decades with treatment morbidities.[7,8] The vast majority of modern OPC patients are treated with RT as a component of their care.[9] Refinements in RT delivery can lessen nontarget doses in hopes of toxicity reduction.[10,11] Yet, even with modern RT, toxicity prediction strategies are needed because collateral dose to normal tissue is unavoidable without compromising the therapeutic RT dose required to eradicate the tumor.[12]

Even with intensity-modulated radiation therapy (IMRT), higher hypothyroidism rates have been reported[13,14] and this phenomenon could be elucidated by the IMRT beam path parameters, namely higher integral radiation dose to nontarget normal tissues.[15] However, with IMRT, application of additional dose constrains that would decrease the delivered RT dose to the thyroid gland (TG) is feasible.[13] To use the modern RT techniques for risk-adapted RT plans, identifying clinical and dosimetric parameters for modeling normal tissue complication probability (NTCP) is an unmet clinical need.[16] Currently, there is a growing effort to identify the relevant input clinical and dosimetric parameters that could be incorporated in an NTCP model for hypothyroidism.[17,18] At present, some dosimetric parameters (ie, mean thyroid gland dose, V30, V40, and V50) have been proposed to minimize the risk of hypothyroidism after RT.[19-22] Nevertheless, NTCP models that account for the patient-intrinsic risk factors[23] and treatment parameters could offer a powerful approach for personalized treatment selection and RT plan optimization based on estimated risk of complications.

This work aims to apply existing NTCP models for radiation-induced hypothyroidism (RHT) risk prediction in a large cohort of OPC patients and to explore the clinical and dosimetric parameters for RHT.

Methods and Materials

Study cohort

We identified 523 OPC patients treated with IMRT without thyroidectomy at University of Texas, MD Anderson Cancer Center, between 2007 to 2013. OPC patients treated with retrievable IMRT plan or dose DICOMs and known baseline thyroid status and available follow-up thyroid function tests were included (Fig E1, available online at https://doi.org/10.1016/j.adro.2019.08.006). In total, 360 OPC patients were eligible for the analysis. Clinical HT that required thyroid replacement therapy was defined as grade ≥2 HT per Common Terminology Criteria for Adverse Events grading system, versions 3 and 4[24,25] at least 6 months post-RT.

Intensity modulate radiation therapy

Treatment planning was conducted using Pinnacle 9.6 software in all patients (Philips Medical Systems, Andover, MA). All patients were treated with IMRT, per our institutional protocol for OPC,[26] with each case undergoing rigorous group peer-review prior treatment

commencement.[27] IMRT was delivered using split-field[28] technique for most of the patients. Whole-field IMRT was used only for bulky tumors, which might be underdosed using the split-field approach. We are following the National Comprehensive Cancer Network guidelines[29] for the sequential and concurrent systemic therapy.

Dosimetric data

Treatment plan and dosimetric data were restored using Pinnacle 14 software (Phillips Medical Systems, Andover, MA). Planning CT DICOM files were exported into a benchmarked[30] commercial deformable registration/segmentation software Velocity AI (Velocity AI 3.0.1, Velocity Medical Solutions, Atlanta, GA). Mean dose (Dmean) to the TG and its volume were calculated using Velocity AI software. Thyroid gland was autosegmented using a previously validated atlas data set[30] and subsequently curated by expert radiation oncologists (MK and ASRM). Dose-volume histograms (DVHs) were extracted from Velocity AI.

Normal tissue complication probability modeling

Two previously published NTCP models for hypothyroidism were selected to be tested on our institution's data set. One model has been published by Boomsma et al,[18] which uses the logistic function.

$$NTCP = \frac{1}{1 + e^{-S}} \quad \text{Equation 1}$$

In Equation 1, S is a linear function determined through a generalized linear model fit based on mean thyroid dose and thyroid volume and has the form.

$$S = 0.011 + (0.062 * mean\ dose\ thyroid\ galnd) + (-0.19 * thyroid\ gland\ volume) \quad \text{Equation 2}$$

The second model was enumerated by Cella et al,[31] and it incorporates thyroid V_{30}, the absolute volume (in mL) of the thyroid gland receiving a dose of 30 Gy, the absolute total volume of the thyroid gland, and sex. The associated S function is Equation 3, which can be applied through Equation 1.

$$S = 1.94 + (0.26 * thyroid\ V_{30}) + (-0.27 * thyroid\ gland\ volume) + (-2.21 * sex\ (male)) \quad \text{Equation 3}$$

For both of the selected models, a model of the same form was fit on our institution's data set similarly using generalized linear model fitting with a logit link function. Additionally, several models similar to Equation 2 but instead using equivalent uniform dose (EUD) were also fit. EUD was calculated according to Equation 4, incorporating values of the a parameter including 0.5, 2, 3, 4, and 5.

$$EUD = \left(\sum_{i=1}^{n} v_i D_i^a \right)^{\frac{1}{a}} \quad \text{Equation 4}$$

where v_i and D_i describe the differential volume and dose, respectively. The a parameter value of 1 was omitted because in that case the EUD is equal to the mean dose.

Model fitting was carried out using the glm function in R version 3.4.2.[32] The predictive ability of the selected models and the fitted models were compared based on receiver operating characteristic (ROC) curves, which were calculated with the pROC package in R.[33] The ROC curves for the fitted models were calculated based on 10-fold repeated cross-validation with 100 repeats. The area under the curve (AUC) values for the various models were then compared according to DeLong's test for 2 correlated ROC curves also using the pROC package.[33]

Statistical analysis

All statistical analysis was performed using commercial statistical analysis software programs (JMP v12 Pro, SAS Institute, Cary, NC; R version 3.4.2, R Foundation for Statistical Computing, Vienna, Austria). Descriptive statistics were calculated. Statistical analyses of the categorical variables were performed using χ^2 tests and t tests for continuous outcomes. Dmean and volume of TG for those with HT were compared with others using Wilcoxon rank-sum test. Univariate regression models were first examined for clinical variables (T-classification, N-classification, sex, age, treatment modalities, RT dose, volume of the RT neck field and primary site). Multivariable models retained confounders that were independently associated ($P < .05$) with the presence of clinical hypothyroidism. A P value of $< .05$ was considered statistically significant. A Bayesian Information Criteria (BIC)-minimizing stepwise forward model was constructed using the significant clinical and dosimetric variables. Patient dose distributions were interrogated via plots of cumulative DVH (using a 1 Gy step [range, 1-75 Gy]) according to presence or absence of HT, with subsequent Wilcoxon rank-sum tests and P values plotted via heat map analysis and Bonferroni correction for multiple comparisons.

Results

The study cohort includes 360 patients who were treated with IMRT for OPC. The median age was 58 years and 340 were men (Table 1). As many as 233 patients

Table 1 Patient and treatment characteristics

Patient and tumor characteristics	All patients	No clinical HT (G0-1)	Clinical HT required medications (G2)	Univariate *P* value
	N = 360	N = 140	N = 233	
	N (%)	N (%)	N (%)	
Dmean mean ± SD, in Gy	47.29 (±10.5)	42.57 (±13.3)	49.86 (±7.4)	<.0001*
TG volume mean ± SD in mL	11.83 (±3.87)	12.77 (±4.49)	11.32 (±3.39)	.0076*
Age median (range), y	58 (33-85)	58 (35-85)	58 (36-82)	.55
Sex				
Male	312 (87)	110 (87)	202 (87)	.98
Female	48 (13)	17 (13)	31 (13)	
Subsite				.0029*
Base of tongue	184 (51)	55 (43)	129 (55)	
Tonsil	164 (46)	71 (56)	93 (40)	
Others	12 (3)	1 (1)	11 (5)	
T				.0182*
1-2	259 (73)	101 (80)	158 (69)	
3-4	97 (27)	25 (20)	72 (31)	
N				.0050*
0	24 (7)	15 (12)	9 (4)	
1-3	333 (93)	111 (88)	231 (96)	
Induction ± CCRT				.0146*
Yes	129 (36)	35 (28)	94 (40)	
No	231 (64)	92 (72)	139 (60)	
CCRT				.0574
Yes	225 (63)	71 (56)	154 (66)	
No	135 (38)	56 (44)	79 (34)	
HPV/P16 status				.554
Positive	279 (78)	102 (80)	177 (76)	
Negative	29 (8)	10 (8)	19 (8)	
Unknown	52 (14)	15 (12)	37 (16)	
Total RT dose, range (Gy)	70 (60-72)	70 (60-70)	70 (60-72)	.054
IMRT				.1654
Split	341 (95)	123 (97)	218 (94)	
Whole field	19 (5)	4 (3)	15 (6)	
Neck irradiation				.0003*
Ipsilateral	45 (13)	27 (21)	18 (8)	
Bilateral	315 (87)	100 (79)	215 (92)	

Abbreviations: Dmean = mean dose; HPV = human papillomavirus; HT = hypothyroidism; IMRT = intensity-modulated radiation therapy; RT = radiation therapy; SD = standard deviation; TG = thyroid gland.

* Significant *P* value; $P \leq .0$.

(65%) developed clinical HT after RT. The median time to develop RHT was 12 months after compilation of RT. Univariate analysis revealed that advanced tumor stage, positive nodal disease, bilateral neck irradiation, receipt of induction chemotherapy, higher Dmean to TG and smaller TG volume were associated with clinical HT. On multivariate analysis Dmean and TG volume remained statistically significant (Table 2). Figure 1 shows the composite DVHs for patients with and without RHT. Our results show that patients with RHT had numerically higher dose delivery across all DVHs than those without. The cumulative DVHs showed a significant separation between the 2 groups. It is interesting that this separation is observed also in low dose regions, hinting that more effort should be spent to minimize the RT dose as low as possible. Even after Bonferroni correction, significant pairwise dose–volume differences were observed. A nonoverlapping confidence interval of dose in 1-Gy bins visually suggests a magnitude difference of $P < .0001$ (denoted in red in the heat map). To account for multiple comparisons and avoid potential error from normal distribution assumptions while illustrating pairwise dose differentials between RHT and no RHT subgroups, a heat map is displayed below to quantify the magnitude of *P* values for each 1-Gy bin (per nonparametric Wilcoxon rank-sum test for each bin).

Two NTCP models for hypothyroidism were fit on our institution's data following from the models selected from the literature. The model based on the Boomsma et al[18] model predictors is represented by Equation 5.

Table 2 Stepwise forward regression model

Model effects	FDR- LogWorth	FDR-*P* value	Whole model *P* value	Whole model ROC, AUC	Whole model BIC
Dmean, median (Gy)	12.326	<.0001	<.0001	0.7299	419.8
TG volume, median (mL)	6.431	<.0001			

Abbreviations: AUC = area under the curve; Dmean = mean dose; FDR = false discovary rate; LogWorth = −log10[*P* value] = such that *P* = .01 is equivalent to a LogWorth of 2.0, *P* = .001 is denoted by LogWorth of 3.0, etc; ROC = receiver operating characteristic; TG = thyroid gland.

$$S = -1.69 + (0.0879 * \textit{mean dose thyroid gland}) + (-0.151 * \textit{thyroid gland volume}) \quad \text{Equation 5}$$

Equation 6 is the model which includes thyroid V_{30}, total thyroid gland volume, and sex as predictors.

$$S = 2.38 + (0.424 * \textit{thyroid } V_{30}) + (-0.557 * \textit{thyroid gland volume}) + (0.253 * \textit{sex}\,(\textit{male})) \quad \text{Equation 6}$$

Table 3 presents a summary of the model parameters for all models, including the parameters for each of the EUD-based models. For each value of a parameter, the corresponding EUD was significantly associated with hypothyroidism on both univariate analysis and multivariable logistic regression.

Table 3 also presents the results for the ROC curve analysis for all models. Each of the models fit on our institution's data are compared with the performance of both the Boomsma et al[18] and Cella et al[31,34] models according to DeLong's test. Model performance was generally similar across all of the models with observed AUC values of approximately 0.7, with the exception of the model from Cella et al[31,34] being slightly lower. The AUC for the Cella et al[31,34] model was 0.66 (95% confidence interval [CI], 0.60-0.72), whereas that for the similar fitted model was 0.71 (95% CI, 0.66-0.77). The difference between the AUCs of the ROC curves for the 2 previously published

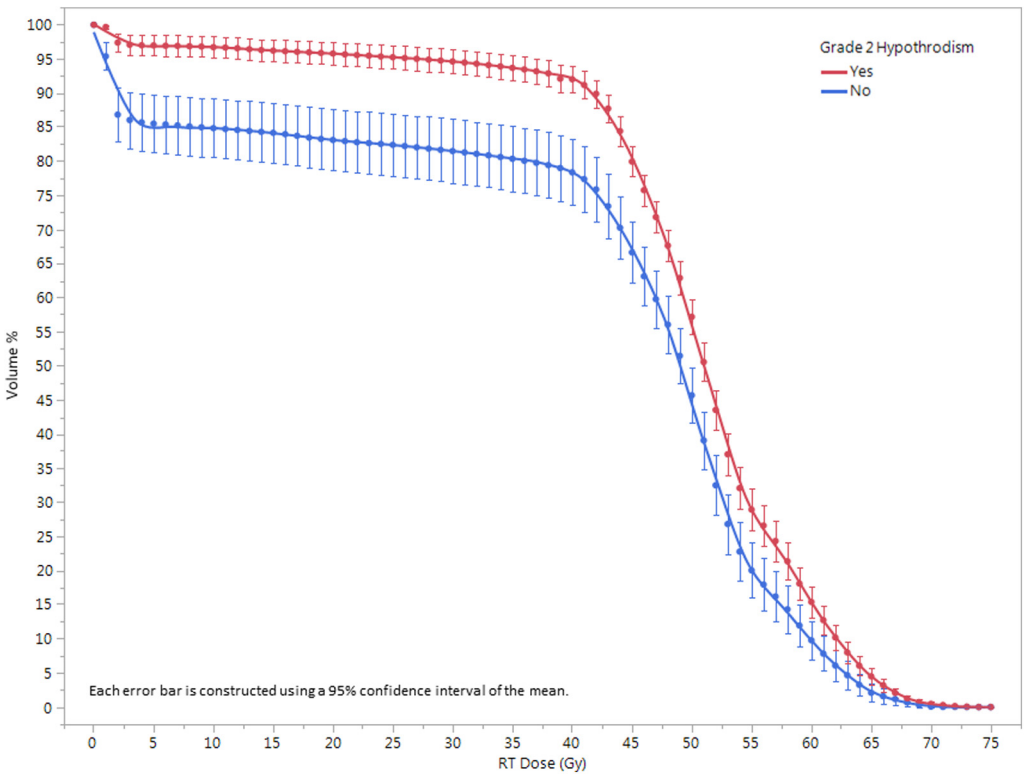

Figure 1 Thyroid gland DVH stratified by radiation therapy-induced clinical hypothyroidism. Comparison of the dose volume histograms (DVHs) between the plans for radiation therapy-induced hypothyroidism cases versus No-HT; each error bar is constructed using a 95% confidence interval of the mean. *P* value thresholding for multiple comparisons was used with $P < .0006$ deemed significance, which indicated in the heat map by the read shading.

Table 3 Summary of model parameters and coefficients for previously published and fitted models

Model source	S equation	Coefficients (SD error)					AUC (95% CI)	DeLong's Test *P* value	
		a	b	c	d			Boomsma et al	Cella et al
Boomsma et al[5]	$a + b * D_{mean} + c * V_{thyroid}$	0.011	0.062	−0.19			0.72 (0.66-0.77)		0.004
Cella et al[31]	$a + b * V_{30} + c * V_{thyroid} + d *$ sex(male)	1.94	0.26 (0.09)	−0.27 (0.11)	−2.21 (0.85)		0.66 (0.60-0.72)	.004	
Fit	$a + b * D_{mean} + c * V_{thyroid}$	−1.49 (0.63)	0.0864 (0.0132)	−0.161 (0.033)			0.72 (0.67-0.78)	.77	.004
Fit	$a + b * V_{30} + c * V_{thyroid} + d *$ sex(male)	2.44 (0.47)	0.416 (0.0731)	−0.559 (0.087)	0.342 (0.384)		0.71 (0.66-0.77)	.39	.013
Fit	$a + b * EUD_{a=0.5} + c * V_{thyroid}$	−0.498 (0.519)	0.0674 (0.0103)	−0.161 (0.033)			0.72 (0.66-0.78)	.78	.005
Fit	$a + b * EUD_{a=2} + c * V_{thyroid}$	−3.47 (0.90)	0.123 (0.019)	−0.157 (0.033)			0.73 (0.67-0.79)	.56	.005
Fit	$a + b * EUD_{a=3} + c * V_{thyroid}$	−4.86 (1.12)	0.146 (0.024)	−0.151 (0.033)			0.72 (0.67-0.78)	.79	.010
Fit	$a + b * EUD_{a=4} + c * V_{thyroid}$	−5.69 (1.28)	0.159 (0.026)	−0.145 (0.033)			0.72 (0.67-0.78)	.96	.017
Fit	$a + b * EUD_{a=5} + c * V_{thyroid}$	−6.14 (1.38)	0.165 (0.028)	−0.140 (0.033)			0.72 (0.66-0.78)	.91	.025

Abbreviations: AUC = area under the curve; CI = confidence interval; SD = standard deviation.
AUC values from ROC curve analysis for each model are presented and AUC values for fitted models are compared with previously published models according to DeLong test for correlated ROC curves.

models as applied to our data set was significant with a *P* value of .004. There was no statistically significant difference between the AUC of the ROC curve for the model from Boomsma et al[18] compared with the AUCs for the models fit on our institution's data. The AUC for the Cella et al[31,34] model was significantly lower ($P < .05$) than the AUC for most of the models fit on our institution's data set. The largest AUC value (0.73; 95% CI, 0.67-0.79) was observed for the EUD-based model with an *a* parameter of 3. AUC values were similar for all tested EUD-based models with an *a* parameter >1.

Figure 2 displays mean dose to the thyroid gland versus thyroid volume for patients with and without hypothyroidism with a linear regression line and 95% confidence interval for each group. Figure 3 provides a visual comparison of NTCP versus mean thyroid dose for several thyroid volumes for the Boomsma et al[18] model and the fitted model of the same form. In addition, 95% confidence bands are provided for the fitted model. These models were selected for comparison because the Boomsma et al model performed better on our internal data set and because they can be directly compared owing to being based on the same input parameters, whereas an EUD based model is not directly comparable to a model based on mean dose. The largest differences between the models are observed for the predictions for thyroid volumes of 25 mL.

Discussion

In a homogenous cohort of 396 OPC patients treated with IMRT, we examined the clinical and dosimetric correlates of RT-induced hypothyroidism and validated existing NTCP models for RT-induced hypothyroidism on a large sample cohort. Our data showed that patients with advanced tumor stage, positive nodal disease, bilateral neck irradiation, receipt induction chemotherapy, higher Dmean to TG, and smaller TG volume had significant risk of clinical HT in the univariate analysis. On multivariate analysis, Dmean and TG volume maintained a statistically significant association with clinical HT. Boomsma et al[5] and Cella et al[31] models for RT-induced hypothyroidism were tested on our institution's data in addition to models fit on the data set. Model performance was generally similar across all of the models with observed AUC values of approximately 0.7, with the exception of the model from Cella et al.[31] Additionally, the corresponding EUD was significantly associated with HT on both univariate analysis and multivariable logistic regression. A linear relationship was observed between the TG volume and Dmean. Within the study cohort, patients with HT were predominately characterized by being female, with average aged, (median [range], 58[36.5-82] years old), diagnosed with BOT cancer, presented with advanced N stages, received induction chemotherapy concurrent chemoradition (IC ± CCRT), had HPV positive status, treated with split

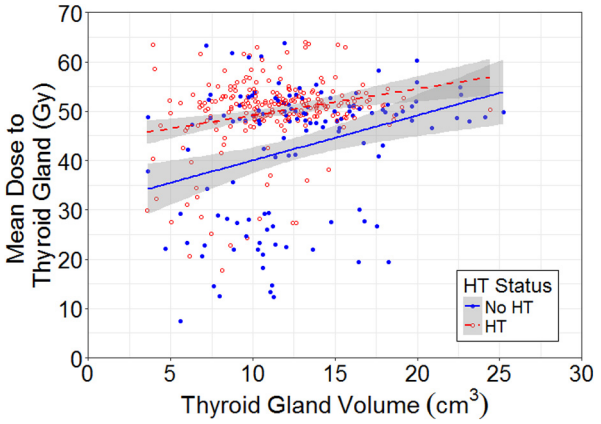

Figure 2 Mean dose to the thyroid gland versus thyroid volume for patients with and without hypothyroidism (HT). A linear regression line and 95% confidence interval are displayed for each group.

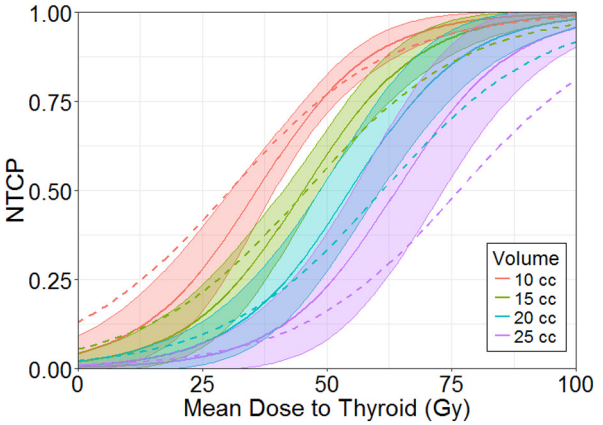

Figure 3 Comparison of predicted NTCP values for different thyroid gland volumes. Comparison of predicted NTCP values for different thyroid gland volumes (colors) for the model previously published by Boomsma et al (dashed lines) and the model fit on our institution's data (solid lines). 95% confidence intervals (bands) are presented for the fitted model.

field IMRT, and underwent bilateral neck irradiation. Patients with HT had smaller TG volume compared with those without (11.8 compared with 12.8 mL, $P < .0001$). Dmean was significantly higher in patients with clinical HT versus those without (50 vs 42 Gy, $P < .0001$).

Differences in the performance between the model from Cella et al[31] and the model from Boomsma et al[18] and the fitted models could potentially be explained by differences in the patient populations on which the models were fitted, different study outcomes, and longer follow-up time. Cella et al[31] developed their model on data from a cohort of patients treated for Hodgkin lymphoma with a median treatment dose of 32 Gy (range, 30-36).[31,35] This is in contrast to our institution's cohort which received a median treatment dose of 69.96 Gy (range, 59.96-72.00). The median thyroid Dmean for the Cella et al[31,35] study was reported for those patients with and without HT separately, with values of 31.5 Gy (range, 30.4-32.6) and 18.9 Gy (range, 15.8-29.8), respectively. However, for our cohort the median thyroid Dmean was 50.23 (range, 7.51-63.92; across all patients), which was greater than the maximum possible Dmean from their study. The Cella et al[31] model was tested on a cohort of breast cancer patients reported by Johansen et al[36] for which median thyroid Dmean was 31 Gy (range, 22-42) for patients with clinical HT (requiring treatment) and 31 Gy (range, 28-28) for those without HT. Cella et al[31] reported AUC values for their model (Equation 3) of 0.874 for their internal data set and 0.914 for the Johansen et al[36] data set. In contrast, the model from Boomsma et al[18,36] (Equation 2) produced AUC values of 0.718 and 0.898,[31] respectively, for the 2 data sets. The Boomsma et al[18] model was fit for a cohort of head and neck cancer patients treated to dose in the range of 46 to 66 Gy; however, the median thyroid Dmean was not reported for this cohort. In this study we evaluated radiation-induced hypothyroidism as clinically evident hypothyroidism of CTCAE grade 2 or higher. This is in contrast to the previous studies which included both subclinical (laboratory determined) and clinical hypothyroidism. Such differences could further contribute to variations in model performance. Given the generally lower AUC values observed in our study (~0.7) across both previously published models and fitted models compared with those AUC values observed in previous publications, it would seem that such models are less suited to predicting clinical hypothyroidism, though the Cella et al[31] model did perform well for the Johansen et al[36] cohort (clinical HT). The differences in thyroid Dmean for the various cohorts, lower values for Cella et al[31] and Johansen et al[36] and higher values for Boomsma et al[18] and our cohort, would also contribute to differences in model performance. It is thus of great importance to continue to develop models based on different patient populations, such that a model whose input data most closely matches the intended application can be selected when evaluating RHT risk.

Our results indicated that NTCP models based on EUD with an a parameter greater than 1 (2, 3, 4, and 5 tested) performed the best, although the model performance was not significantly better than that of the Boomsma et al model. There may be some value in further investigation into the role of EUD in the NTCP for hypothyroidism; however, this is beyond the scope of this study.

Hypothyroidism is frequently observed after radiation[1-5] and this phenomenon could be explained by RT-induced direct cell injury, microvascular insult[1,37] and immune- mediated damage,[38] which resulted in reduction of the TG volume.[39] Such volumetric reduction was found to be correlated with the Dmean to the TG rather than the Dmax, which only affect a relative small volume of the TG.[40] A thyroid volume effect in RHT development was found in our current research effort and others[13,36,41]; there is a decrease in the risk of RHT with

larger thyroid gland volumes and the volume of irradiated thyroid seems to be a risk factor for RHT.[3]

Our study is the largest of its kind investigating the risk of RHT after IMRT for HNC patients with a comprehensive approach incorporating the clinical and dosimetric characteristics in the risk assessment. The study cohort is a homogenous sample of OPC patients treated with definitive intent and a median IMRT dose of 70 Gy. Our methodology accounts for the EUD in comparison to other approaches. However, our study has the caveats of retrospective design and utilization of single institution data set.

Multi-institutional collaboration and prospective data are incredibly needed to construct and validate the novel predictive models for treatment-induced toxicities. Although we excluded patients with unknown thyroid status at pre-RT time point, the retrospective nature of the study did not allow for accounting for the possibility of missing pre-RT subclinical hypothyroidism status. This caveat may introduce some bias. Prospective studies with frequent monitoring of the TSH level at multiple time points (baseline, during RT and shortly after RT (during the acute or subacute phase [up to 6 months after RT]) are needed to account for the fluctuation nature of the TSH results, false laboratory results and possibility of spontaneous recovery of pre-RT and acute or subacute subclinical hypothyroidism, which may affect the accuracy of the assessment of the thyroid status later and identify patients with late RT-induced hypothyroidism.

Our data confirms that thyroid volume and mean thyroid radiation therapy dose are important predictors of clinical radiation therapy-induced hypothyroidism. Accordingly, personalized plan optimization, based on individual thyroid volume, is recommended to reduce the risk of clinically relevant hypothyroidism after IMRT for OPC. In the era of the HPV-driven OPC, rapidly growing numbers of young survivors live longer with treatment morbidities, thus, maintaining the functional outcome is the metric of therapeutic success.

Supplementary data

Supplementary material for this article can be found at https://doi.org/10.1016/j.adro.2019.08.006.

References

1. Cannon CR. Hypothyroidism in head and neck cancer patients: Experimental and clinical observations. *Laryngoscope*. 1994;104:1-21.
2. DeGroot LJ. Radiation and thyroid disease. *Bailliere's clinical endocrinology and metabolism*. 1988;2:777-791.
3. Alterio D, Jereczek-Fossa BA, Franchi B, et al. Thyroid disorders in patients treated with radiotherapy for head-and-neck cancer: A retrospective analysis of seventy-three patients. *Int J Radiat Oncol Biol Phys*. 2007;67:144-150.
4. Mercado G, Adelstein DJ, Saxton JP, Secic M, Larto MA, Lavertu P. Hypothyroidism: A frequent event after radiotherapy and after radiotherapy with chemotherapy for patients with head and neck carcinoma. *Cancer*. 2001;92:2892-2897.
5. Boomsma MJ, Bijl HP, Langendijk JA. Radiation-induced hypothyroidism in head and neck cancer patients: A systematic review. *Radiother Oncol*. 2011;99:1-5.
6. Vigário P, Teixeira P, Reuters V, et al. Perceived health status of women with overt and subclinical hypothyroidism. *Med Princ Pract*. 2009;18:317-322.
7. Licitra L, Perrone F, Bossi P, et al. High-risk human papillomavirus affects prognosis in patients with surgically treated oropharyngeal squamous cell carcinoma. *J Clin Oncol*. 2006;24:5630-5636.
8. Fakhry C, Westra WH, Li S, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. *J Natl Cancer Inst*. 2008;100:261-269.
9. Adelstein DJ, Ridge JA, Brizel DM, et al. Transoral resection of pharyngeal cancer: Summary of a National cancer Institute head and neck cancer Steering Committee clinical trials planning Meeting, November 6-7, 2011. Arlington, Virginia. *Head Neck*. 2012;34:1681-1703.
10. Head MDA, Neck Cancer Symptom Working G. Beyond mean pharyngeal constrictor dose for beam path toxicity in non-target swallowing muscles: Dose-volume correlates of chronic radiation-associated dysphagia (RAD) after oropharyngeal intensity modulated radiotherapy. *Radiother Oncol*. 2016;118:304-314.
11. van der Molen L, Heemsbergen WD, de Jong R, et al. Dysphagia and trismus after concomitant chemo-intensity-modulated radiation therapy (chemo-IMRT) in advanced head and neck cancer; dose–effect relationships for swallowing and mastication structures. *Radiother Oncol*. 2013;106:364-369.
12. Garden AS, Morrison WH, Wong PF, et al. Disease-control rates following intensity-modulated radiation therapy for small primary oropharyngeal carcinoma. *Int J Radiat Oncol Biol Phys*. 2007;67:438-444.
13. Diaz R, Jaboin JJ, Morales-Paliza M, et al. Hypothyroidism as a consequence of intensity-modulated radiotherapy with concurrent taxane-based chemotherapy for locally advanced head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. 2010;77:468-476.
14. Murthy V, Narang K, Ghosh-Laskar S, Gupta T, Budrukkar A, Agrawal JP. Hypothyroidism after 3-dimensional conformal radiotherapy and intensity-modulated radiotherapy for head and neck cancers: Prospective data from 2 randomized controlled trials. *Head Neck*. 2014;36:1573-1580.
15. Rosenthal DI, Chambers MS, Fuller CD, et al. Beam path toxicities to non-target structures during intensity-modulated radiation therapy for head and neck cancer. *Int J Radiat Oncol Biol Phys*. 2008;72:747-755.
16. Trott K-R, Doerr W, Facoetti A, et al. Biological mechanisms of normal tissue damage: Importance for the design of NTCP models. *Radiother Oncol*. 2012;105:79-85.
17. Emami B, Lyman J, Brown A, et al. Tolerance of normal tissue to therapeutic irradiation. *Int J Radiat Oncol Biol Phys*. 1991;21:109-122.
18. Boomsma MJ, Bijl HP, Christianen MEMC, et al. A prospective cohort study on radiation-induced hypothyroidism: Development of an NTCP model. *Int J Radiat Oncol Biol Phys*. 2012;84:e351-e356.
19. Lee V, Chan S-Y, Choi C-W, et al. Dosimetric predictors of hypothyroidism after radical intensity-modulated radiation therapy for non-metastatic nasopharyngeal carcinoma. *Clin Oncol (R Coll Radiol)*. 2016;28:e52-e60.
20. Zhai R-p, Kong F-f, Du C-r, Hu C-s, Ying H-m. Radiation-induced hypothyroidism after IMRT for nasopharyngeal carcinoma: Clinical and dosimetric predictors in a prospective cohort study. *Oral Oncology*. 2017;68:44-49.

21. Akgun Z, Atasoy BM, Ozen Z, et al. V30 as a predictor for radiation-induced hypothyroidism: A dosimetric analysis in patients who received radiotherapy to the neck. *Radiother Oncol*. 2014;9:104.
22. Chyan A, Chen J, Shugard E, Lambert L, Quivey JM, Yom SS. Dosimetric predictors of hypothyroidism in oropharyngeal cancer patients treated with intensity-modulated radiation therapy. *Radiother Oncol*. 2014;9:269.
23. Vogelius IR, Bentzen SM, Maraldo MV, Petersen PM, Specht L. Risk factors for radiation-induced hypothyroidism. *Cancer*. 2011;117:5250-5260.
24. US Department of Health and Human Services-National Institutes of Health NCI. *Common Terminology Criteria for Adverse Events (CTCAE)*. version 4.0. 2009.
25. Common Terminology Criteria for Adverse Events v3.0 (CTCAE). August 9, 2006. Available from: https://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/ctcaev3.pdf. Accessed June 15, 2018.
26. Garden AS, Kies MS, Morrison WH, et al. Outcomes and patterns of care of patients with locally advanced oropharyngeal carcinoma treated in the early 21(st) century. *Radiat Oncol*. 2013;8:21.
27. Cardenas CE, Mohamed ASR, Tao R, 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;98:532-540.
28. Dabaja B, Salehpour MR, Rosen I, et al. Intensity-modulated radiation therapy (IMRT) of cancers of the head and neck: Comparison of split-field and whole-field techniques. *Int J Radiat Oncol Biol Phys*. 2005;63:1000-1005.
29. NCCN Guidlines version 1.2018 head and neck cancer: National comprehensive cancer Network. Available from: https://www.nccn.org/professionals/physician_gls/pdf/head-and-neck.pdf. Accessed June 15, 2018.
30. Mohamed ASR, Ruangskul M-N, Awan MJ, et al. Quality assurance assessment of diagnostic and radiation therapy-simulation CT image registration for head and neck radiation therapy: Anatomic region of interest-based comparison of rigid and deformable algorithms. *Radiol*. 2015;274:752-763.
31. Cella L, Liuzzi R, Conson M, D'Avino V, Salvatore M, Pacelli R. Development of multivariate NTCP models for radiation-induced hypothyroidism: A comparative analysis. *Radiat Oncol*. 2012;7:224.
32. Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.
33. Robin X, Turck N, Hainard A, et al. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77.
34. Cella L, Conson M, Liuzzi R, Pacelli R. In regard to Boomsma et al. *Int J Radiat Oncol Biol Phys*. 2013;85:11.
35. Cella L, Conson M, Caterino M, et al. Thyroid V30 predicts radiation-induced hypothyroidism in patients treated with sequential chemo-radiotherapy for Hodgkin's lymphoma. *Int J Radiat Oncol Biol Phys*. 2012;82:1802-1808.
36. Johansen S, Reinertsen K, Knutstad K, Olsen D, Fosså S. Dose distribution in the thyroid gland following radiation therapy of breast cancer-a retrospective study. *Radiat Oncol*. 2011;6:68.
37. Eckert CT, Probstein JG, Galinson S. Radiation of the thyroid. *Radiol*. 1937;29:40-44.
38. Cutuli B, Quetin P, Rodier J-F, Barakat P, Grob J-C. Severe hypothyroidism after chemotherapy and locoregional irradiation for breast cancer. *Radiat Oncol*. 2000;57:103-105.
39. Cheng SCH, Wu VWC, Kwong DLW, et al. Sonographic appearance of thyroid glands in patients treated with intensity-modulated radiotherapy or conventional radiotherapy for nasopharyngeal carcinoma. *J Clin Ultrasound*. 2015;43:210-223.
40. Lin Z, Wu VW-C, Lin J, Feng H, Chen L. A longitudinal study on the radiation-induced thyroid gland changes after external beam radiotherapy of nasopharyngeal carcinoma. *Thyroid*. 2011;21:19-23.
41. Kumpulainen EJ, Hirvikoski PP, Virtaniemi JA, et al. Hypothyroidism after radiotherapy for laryngeal cancer. *Radiother Oncol*. 2000;57:97-101.