

Scientific Article

Evaluation of classification and regression tree (CART) model in weight loss prediction following head and neck cancer radiation therapy

Zhi Cheng MD MPH ^a, Minoru Nakatsugawa PhD ^{a,b}, Chen Hu PhD ^c, Scott P. Robertson PhD ^a, Xuan Hui MD MS ^a, Joseph A. Moore PhD ^a, Michael R. Bowers BS ^a, Ana P. Kiess MD PhD ^a, Brandi R. Page MD ^a, Laura Burns BSN ^a, Mariah Muse BSN ^a, Amanda Choflet MS RN OCN ^a, Kousuke Sakaue MS ^d, Shinya Sugiyama MS ^d, Kazuki Utsunomiya MS ^d, John W. Wong PhD ^a, Todd R. McNutt PhD ^a, Harry Quon MD MS ^{a,*}

^a Department of Radiation Oncology and Molecular Radiation Sciences, Johns Hopkins University, Baltimore, Maryland

^b Toshiba America Research, Inc., Baltimore, Maryland

^c Oncology Center—Biostatistics/Bioinformatics, Johns Hopkins University, Baltimore, Maryland

^d Canon Medical Systems Corporation, Otawara, Japan

Received 2 June 2017; received in revised form 2 October 2017; accepted 30 November 2017

Abstract

Objective: We explore whether a knowledge–discovery approach building a Classification and Regression Tree (CART) prediction model for weight loss (WL) in head and neck cancer (HNC) patients treated with radiation therapy (RT) is feasible.

Methods and materials: HNC patients from 2007 to 2015 were identified from a prospectively collected database Oncospace. Two prediction models at different time points were developed to predict weight loss ≥ 5 kg at 3 months post-RT by CART algorithm: (1) during RT planning using patient demographic, delineated dose data, planning target volume–organs at risk shape relationships data and (2) at the end of treatment (EOT) using additional on-treatment toxicities and quality of life data.

Results: Among 391 patients identified, WL predictors during RT planning were International Classification of Diseases diagnosis; dose to masticatory and superior constrictor muscles, larynx, and

Sources of support: Funding for this research was provided by Canon Medical Systems Corporation, the Commonwealth Foundation, Elekta, Philips Radiation Oncology Systems, and the Johns Hopkins University.

Conflicts of interest: M.N. is an employee of Toshiba America Research and was involved with study design, analysis, and writing of the manuscript and reports personal fees from Toshiba America Research, Inc. K.S., S.S., and K.U. report personal fees from Toshiba Medical Systems Corporation during the conduct of the study. J.W. reports grants from Toshiba Medical Systems, grants and other from Elekta, other from JPLC associates, and grants from the National Cancer Institute, outside the submitted work.

* Corresponding author. Department of Radiation Oncology and Molecular Radiation Sciences, Johns Hopkins University, 401 North Broadway, Suite 1440, Baltimore, MD 21231.

E-mail address: hquon2@jhmi.edu (H. Quon).

<https://doi.org/10.1016/j.adro.2017.11.006>

2452-1094/© 2018 Published by Elsevier Inc. on behalf of the American Society for Radiation Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

parotid; and age. At EOT, patient-reported oral intake, diagnosis, N stage, nausea, pain, dose to larynx, parotid, and low-dose planning target volume–larynx distance were significant predictive factors. The area under the curve during RT and EOT was 0.773 and 0.821, respectively.

Conclusions: We demonstrate the feasibility and potential value of an informatics infrastructure that has facilitated insight into the prediction of WL using the CART algorithm. The prediction accuracy significantly improved with the inclusion of additional treatment-related data and has the potential to be leveraged as a strategy to develop a learning health system.

© 2018 Published by Elsevier Inc. on behalf of the American Society for Radiation Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Precision care is an emerging approach for the prevention and treatment of disease states that accounts for individual heterogeneity in genes, environment, and life-style.¹⁻⁸ To accomplish the goal of precision radiation medicine, our group has leveraged an informatic infrastructure centered around our Oncospace database,^{9,10} which contains data captured at the point of care (PoC) during the routine clinical care of cancer patients (Fig 1). Our overarching hypothesis is that this diverse and comprehensive clinical dataset growing within our informatics infrastructure can facilitate the requisite knowledge-discovery and development of machine learning-based predictive algorithms that offer particular advantages compared with traditional approaches¹¹ to guide individualized

patient decisions as a component of a learning health system.¹²⁻¹⁴

Weight loss is a common side effect among head and neck cancer (HNC) patients throughout treatment and follow-up¹⁵ (Fig 2) and results from multiple toxicities. During chemotherapy radiation therapy (CRT), critical weight loss has been reported to be as high as 57%¹⁶ in irradiated HNC patients. HNC patients are at risk of reduced oral intake because of CRT-induced side effects such as taste and/or smell alternations, painful mucosal inflammation of the oral cavity and pharynx, secretion changes (xerostomia or thick, ropey secretions), and swallow muscle damage caused by radiation injuries and/or surgical ablation of muscular and nervous structures.¹⁷ Moreover, these symptoms may become progressively worse during and after treatment, compromising quality of life (QoL) and contributing

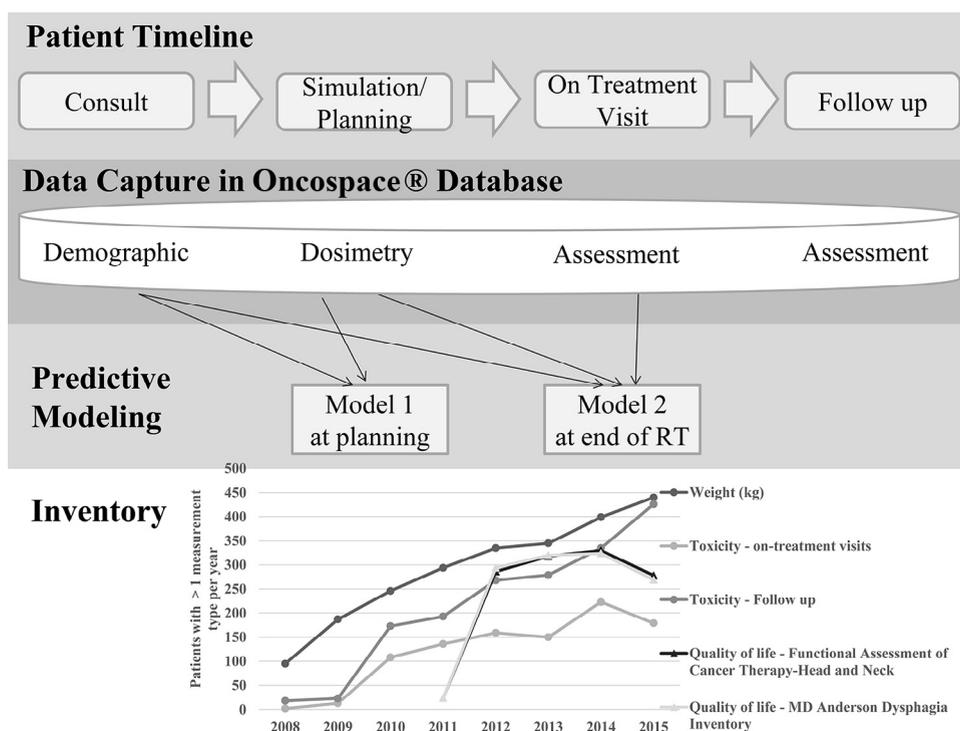


Figure 1 Timeline for patient care, data capture, and storage in the Oncospace database and weight loss prediction models (top). Data inventory with weight, toxicities, and patient-reported quality of life among head and neck cancer patients from 2008 through 2015 (bottom). RT, radiation therapy.

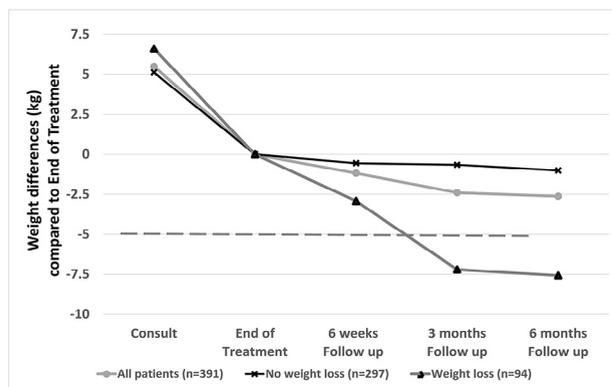


Figure 2 Trend of mean weight difference from the end of treatment (kilograms). Patients with and without critical weight loss (≥ 5 kg) at 3 months posttreatment are stratified.

to poor treatment outcomes and prognosis.^{16,18,19} It is important for care providers to identify patients at risk of critical weight loss, with the goal of effective weight management or, ideally, the prophylactic prevention of malnutrition. Interventions include possible treatment plan modifications, feeding tube placement, pain management, swallow exercise, and frequent monitoring, which can be used during treatment or posttreatment follow-up if the patient is at risk for critical weight loss after treatment.

The causal relationship between RT and weight loss is highly multifactorial. To identify patients at risk of critical weight loss, we applied an exploratory knowledge-discovery approach. Several studies have investigated the predictive risk factors of weight loss during or after the treatment, mainly using clinical factors.^{20,21} In this study, we used all the dosimetric information available in an RT plan in addition to the clinical factors and piloted the use of the classification and regression tree (CART) algorithm to develop a predictive algorithm for critical weight loss in irradiated HNC patients.

Methods and materials

Database and data collection

Oncospace is an analytical database that has systematically aggregated prospectively collected clinical outcomes, RT plans, and geometric shape relationship at our institution since 2007 (Fig 1).^{6,9} Clinical data such as patient demographic information, medical histories, physician-assessed toxicities, and patient-reported QoL were directly and routinely collected into the database at the PoC, facilitated by a Web interface to the database (Fig 1). Patients were treated with intensity modulated RT (IMRT) for 6 to 7 weeks with either a once- or twice-daily hyperfractionated schedule. Toxicity grading defined by the Common Terminology Criteria for Adverse Events (versions 3.0 and 4.0)

was assessed during weekly on-treatment visits and follow-up visits; patient-reported outcomes (PROs) including QoL were measured by the Functional Assessment of Cancer Therapy-Head and Neck, and the MD Anderson Dysphagia Inventory instruments²²⁻²⁶ were assessed at the end of treatment (EOT) and follow-up visits. For RT, prescribed dose to the primary tumor volume and clinically involved nodes was 70 to 72 Gy with 2.0 to 2.2 Gy per fraction, combined with standard dose to moderate- (63 Gy) and low-risk (57 Gy) nodal volumes. Comprehensive dosimetric data defined as the dose delivered at 1% volume increments for each high-/mid-/low-risk planning target volume (PTV) and each organ at risk (OAR) were computed in Oncospace.⁹

Study inclusion criteria were HNC patients treated with IMRT \pm concurrent chemotherapy \pm induction chemotherapy from 2007 through 2015 and patients with weight measurements both at the end of treatment and at 3 month follow-up. Patients with multiple treatments, reirradiation and palliative RT, and missing weight measurements at these time points were excluded.

Statistical design and analysis

Weight loss of ≥ 5 kg (yes/no) at 3 months (60-120 days) posttreatment from the EOT was the primary study endpoint (Fig 2). To potentiate decision support through the RT, 2 prediction models were developed: model 1 at the time of the RT planning with variables available before RT and model 2 at the EOT with additional clinical assessments variables captured at EOT. Model 1 used diagnosis as coded by the International Statistical Classification of Diseases and Related Health Problems (ICD-9) code, TNM staging, age, and dose-volume histogram data at 1% volume increments for OAR as follows: (1) muscles: inferior/middle/superior constrictors, cricopharyngeus, and masticatory (masticator, temporalis, medial, and lateral pterygoid); (2) swallow pathway: oral cavity, oral mucosa, soft palate, and larynx; and (3) secretion: parotid and submandibular glands. Lateralized OARs (eg, left and right parotid glands) were combined into a single structure for calculating dose-volume histogram. Minimum distances (centimeters) between high-/mid-/low-risk PTV and OARs were extracted from the records of geometric shape relationships calculated by the overlap volume histogram algorithm.²⁷⁻³⁰ Model 2 included additional data such as toxicity and QoL. For knowledge-discovery purposes, QoL questions were treated as individual information and analyzed independently.

Baseline characteristics were compared with binary weight loss status (≥ 5 vs < 5 kg) at 3 months posttreatment by *t* test and Wilcoxon test for continuous variables and χ^2 test for categorical variables (Table 1). Statistical significance was defined as $P < .05$.

Decision tree models predicting weight loss were built with the CART algorithm.³¹ The cohort was partitioned into

Table 1 Study population characteristics for patients by critical weight loss (<5 vs ≥5 kg) at RT planning and at the end of RT

| Parameters | N | No CWL reference | CWL | P value |
|---------------------------------------|---------------|------------------|---------------|---------|
| At RT planning: demographic | N = 391 | N = 297 | N = 94 | |
| Age, mean (SD) | 57.39 (10.69) | 56.92 (10.99) | 58.84 (9.76) | .133 |
| Sex, n (%) | | | | .901 |
| Man | 306 (78.26) | 232 (78.11) | 74 (78.72) | |
| Women | 85 (21.74) | 65 (21.89) | 20 (21.28) | |
| Race, n (%) | | | | .111 |
| Caucasian | 187 (76.64) | 128 (74.42) | 59 (81.94) | |
| African American | 42 (17.21) | 32 (18.60) | 10 (13.89) | |
| Asian | 4 (1.64) | 4 (2.33) | 0 (0.00) | |
| Others | 10 (4.10) | 8 (4.65) | 2 (1.39) | |
| Tumor site, n (%) | | | | .040 |
| Oral cavity | 3 (0.77) | 2 (0.67) | 1 (1.06) | |
| Nasopharynx | 14 (3.58) | 9 (3.03) | 5 (5.32) | |
| Oropharynx | 157 (40.15) | 108 (36.36) | 49 (52.13) | |
| Hypopharynx | 9 (2.30) | 7 (2.36) | 2 (2.13) | |
| Larynx | 46 (11.76) | 41 (13.80) | 5 (5.32) | |
| Other | 162 (41.43) | 130 (43.77) | 32 (34.04) | |
| Tumor classification, n (%) | | | | .174 |
| 0 | 21 (7.22) | 12 (5.45) | 9 (12.68) | |
| 1 | 60 (20.62) | 47 (21.36) | 13 (18.31) | |
| 2 | 96 (32.99) | 72 (32.73) | 24 (33.80) | |
| 3 | 41 (14.09) | 35 (15.91) | 6 (8.45) | |
| 4 | 73 (25.09) | 54 (24.55) | 19 (26.76) | |
| N classification, n (%) | | | | .494 |
| 0 | 77 (26.55) | 62 (28.31) | 15 (21.13) | |
| 1 | 44 (15.17) | 31 (14.16) | 13 (18.31) | |
| 2 | 161 (55.52) | 121 (55.25) | 40 (56.34) | |
| 3 | 8 (2.76) | 5 (2.28) | 3 (4.23) | |
| M classification, n (%) | | | | .073 |
| No | 255 (98.84) | 197 (99.49) | 58 (96.67) | |
| Yes | 3 (1.16) | 1 (0.51) | 2 (3.33) | |
| Overall stage, n (%) | | | | .707 |
| 0 | 3 (1.22) | 2 (1.06) | 1 (1.75) | |
| 1 | 17 (6.94) | 15 (7.98) | 2 (3.51) | |
| 2 | 19 (7.76) | 14 (7.45) | 5 (8.77) | |
| 3 | 37 (15.10) | 30 (15.96) | 7 (12.28) | |
| 4 | 169 (68.98) | 127 (67.55) | 127 (73.68) | |
| Treatment modality, n (%) | | | | .240 |
| RT alone | 127 (32.48) | 99 (33.33) | 28 (29.79) | |
| CRT | 253 (64.71) | 191 (64.31) | 62 (65.96) | |
| CRT + surgery | 8 (2.05) | 4 (1.35) | 4 (4.26) | |
| Surgery + RT | 3 (0.77) | 3 (1.01) | 0 (0.00) | |
| Smoking status, n (%) | | | | .807 |
| Never smoked | 67 (40.36) | 58 (41.13) | 9 (36.00) | |
| Quit smoking | 90 (54.22) | 75 (53.19) | 15 (60.00) | |
| Currently smoking | 9 (5.42) | 8 (5.67) | 1 (4.00) | |
| HPV, n (%) | | | | .810 |
| Yes | 56 (58.95) | 47 (59.49) | 9 (56.25) | |
| No | 39 (41.05) | 32 (40.51) | 7 (43.75) | |
| At RT planning: dosimetric | | | | |
| PTV D95 dose, n (%) | | | | .193 |
| ≤65 Gy | 89 (23.73) | 72 (25.35) | 17 (18.68) | |
| >65 Gy | 286 (76.27) | 212 (74.65) | 74 (81.31) | |
| Combined parotid D95, mean (SD) | 287 | 8.61 (5.89) | 10.85 (6.14) | .002 |
| Combined parotid D50, mean (SD) | 287 | 25.61 (13.45) | 28.98 (12.23) | .021 |
| Combined submandibular D95, mean (SD) | 232 | 42.25 (19.97) | 47.77 (17.06) | .044 |

(continued on next page)

Table 1 (continued)

| Parameters | N | No CWL reference | CWL | P value |
|----------------------------------------------|-------------|------------------|---------------|---------|
| Combined submandibular D50, mean (SD) | 232 | 54.08 (19.85) | 58.51 (15.92) | .206 |
| Larynx D95, mean (SD) | 258 | 27.56 (17.2) | 31.37 (16.21) | .248 |
| Larynx D50, mean (SD) | 258 | 37.61 (17.05) | 43.21 (14.92) | .339 |
| Combined masticatory muscle D95, mean (SD) | 150 | 11.98 (10.7) | 17.39 (10.22) | .002 |
| Combined masticatory muscle D50, mean (SD) | 150 | 33.71 (16.53) | 41.5 (13.67) | .002 |
| Superior constrictor muscle D95, mean (SD) | 149 | 47.81 (17.21) | 53.03 (14.35) | .046 |
| Superior constrictor muscle D50, mean (SD) | 149 | 56.66 (14.88) | 60.05 (13.82) | .044 |
| Middle constrictor muscle D95, mean (SD) | 142 | 45.32 (16.36) | 47.84 (12.67) | .634 |
| Middle constrictor muscle D50, mean (SD) | 142 | 53.65 (15.07) | 56.35 (12.23) | .397 |
| Inferior constrictor muscle D95, mean (SD) | 141 | 36.78 (16.15) | 40.45 (12.6) | .188 |
| Inferior constrictor muscle D50, mean (SD) | 141 | 43.91 (15.99) | 46.36 (12.24) | .460 |
| Pharyngeal constrictor muscle D95, mean (SD) | 75 | 22.32 (15.73) | 24.13 (15.44) | .631 |
| Pharyngeal constrictor muscle D50, mean (SD) | 75 | 40.04 (16.65) | 35.18 (20.41) | .596 |
| Cricopharyngeal muscle D95, mean (SD) | 107 | 34.27 (14.67) | 40.93 (10.77) | .043 |
| Cricopharyngeal muscle D50, mean (SD) | 107 | 38.34 (15.48) | 44.27 (10.3) | .121 |
| At the end of RT: additional assessments | | | | |
| Karnofsky performance status, n (%) | | | | |
| ≥80 | 282 (74.80) | 215 (75.44) | 67 (72.83) | .616 |
| <80 | 95 (25.20) | 70 (24.56) | 25 (27.17) | |
| Pain at EOT, n (%) | | | | |
| ≥5 | 109 (27.88) | 79 (26.00) | 30 (31.91) | .317 |
| <5 | 282 (72.12) | 218 (73.40) | 64 (68.09) | |
| Dysphagia at EOT, mean (SD) | 332 | 0.72 (0.99) | 0.85 (0.98) | .318 |
| Xerostomia at EOT, mean (SD) | 391 | 1.41 (0.75) | 1.59 (0.63) | .030 |
| Food oral intake score, n (%) | | | | |
| ≥4 | 113 (84.33) | 99 (86.84) | 14 (70.00) | .056 |
| <4 | 21 (15.67) | 15 (13.16) | 6 (30.00) | |
| Penetration Aspiration Scale, n (%) | | | | |
| ≤2 | 66 (70.21) | 52 (69.33) | 14 (73.68) | .711 |
| >2 | 28 (29.79) | 23 (30.67) | 5 (26.32) | |
| PEG/NG tube used, n (%) | | | | |
| Yes | 122 (75.31) | 86 (72.27) | 36 (83.72) | .134 |
| No | 40 (24.69) | 33 (27.73) | 7 (17.50) | |

CRT, chemotherapy radiation therapy; CWL, critical weight loss; D95, dose to 95% of muscle; EOT, end of treatment; HPV, human papilloma virus; NG, nasogastric; PEG, percutaneous endoscopic gastrostomy; PTV, planning target volume; RT, radiation therapy; SD, standard deviation.

2 subsets based on a predictor variable with the highest risk of weight loss. This process was repeated on each derived subset in an iterative manner (ie, recursive partitioning). At each partition, the information gain criteria were used to determine the partitioning predictor variable with a binary threshold for continuous or ordinal variables and with binary partitioning.

To handle the missing values in each primary predictor, surrogate variables were used to split the observations missing the primary predictor. A maximum of 5 surrogates were found for each primary predictor by reapplying the partitioning algorithm (without recursion) to estimate the binary split of the primary predictor. If all 5 surrogates were missing, then the observation was sent in the majority direction.^{32,33}

Leave-one-out cross-validation was used to determine the number of predictor variables selected in the decision tree models and to calculate the model performance. The DeLong test was further used to compare the area under the curve

of the 2 models. All the statistical analysis were performed by R 3.2.2 with the rpart package.³⁴

Results

A total of 391 patients met the study inclusion criteria. The study population consisted of more men (n = 306, 78%) and early stage (T ≤ 2) cancer (n = 187, 75%). The mean age was 57 years, and the mean weight loss at 3 months posttreatment was 2.4 kg. Nearly one-half of the study population had tumors located in the pharynx (nasopharynx, oropharynx, and hypopharynx) (n = 180, 46%) (Table 1).

Predictors for critical weight loss

Patient lost weight throughout radiation treatment. Mean weight did not decrease further after the end of the treatment,

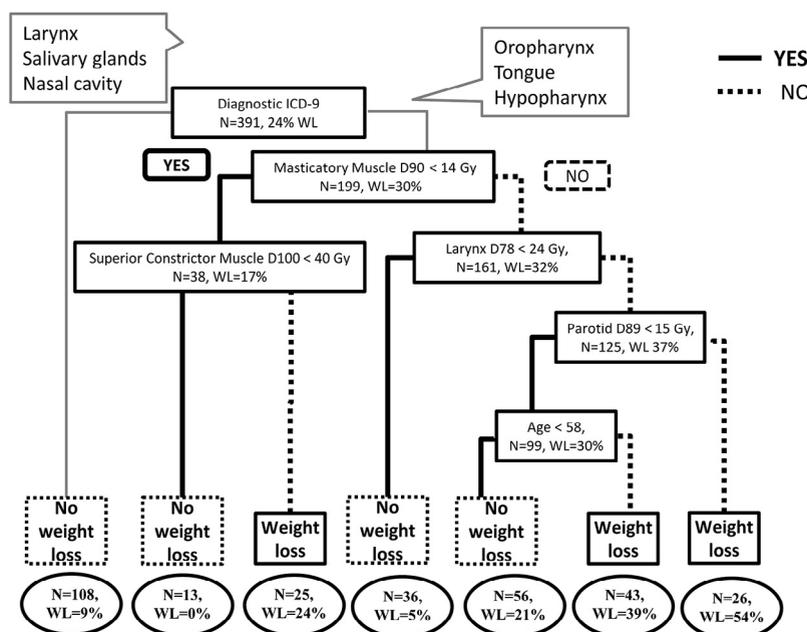


Figure 3 Prediction tree during planning phase (model 1) for critical (≥ 5 kg) WL among head and neck cancer patients 3 months posttreatment. Nodes display sample size (n) and the percentage of patients with critical weight loss within the (sub)group. The potential risk factors in the CART analysis were ICD-9 code; dose to masticatory muscle, superior constrictor muscle, larynx, and parotid; and age. CART, classification and regression tree; D100, dose to 100% of organ; ICD-9, International Classification of Diseases-9; WL, weight loss.

and started to increase at the 3-month follow-up (Fig 2). At the 3-month follow-up, 94 (24%) patients continued to lose weight beyond the 5 kg threshold. Analyses using a 5% weight loss threshold are presented in eTable 1 (available as supplementary material online only at www.practical.radonc.org).

Model 1 was developed at the planning phase of RT when dosimetric and tumor-related variables were available (Fig 3). Anatomic tumor site by ICD-9 code was particularly important. Dose to 90% of the masticatory muscle (D90) with threshold of 14 Gy, D100 (40 Gy) of superior constrictor muscle, D78 (24 Gy) of larynx, and D89 (15 Gy) of parotid glands were also selected from RT and dosimetric variables. Age (58 years) was the last predictor of critical weight loss. The area under the curve, sensitivity, specificity, and positive predictive and negative predictive values by cross-validation were 0.77, 0.77, 0.67, 0.43, and 0.90, respectively (Table 2).

To interpret model 1, the principal discriminating node was tumor site (ICD-9 code), which is consistent with previous reports.³⁵⁻³⁷ Patients with laryngeal tumors were at a lower risk of weight loss compared with patients with pharyngeal tumors. This was consistent with the univariate analysis (Table 1; tumor site, $P = .04$) and suggests that the propulsive swallowing structures may be more important than the coordinating role of the larynx in maintaining weight. Among the pharyngeal cancer patients, the next partition was the radiation dose to the masticatory muscles, which are responsible for chewing, and again underscore the effect

of irradiating muscles involved in the swallow process. Patients with low doses to the masticatory muscles were further partitioned by the dose to the superior constrictor muscle, which was the most superiorly located muscle of the 3 pharyngeal constrictors that moved boluses downward into the esophagus. Patients who received >14 Gy to the masticatory muscles were further split by D78 larynx. For patients who received a larynx dose of ≥ 24 Gy, parotid D89 and age (ie, greater or less than 15 Gy and 58 years of age) were the last determinants of significant weight loss; parotid D89 and age affect swallow and parotid function, leading to weight changes. These findings were consistent with previous studies³⁸ suggesting xerostomia was associated with a threshold of 26 Gy; late dysphagia was associated with a mean dose to pharyngeal constrictors and larynx ≤ 50 Gy.

The prediction tree at EOT (model 2) included the additional clinical assessment variables during treatment

Table 2 Characteristics of the weight loss prediction models (<5 vs ≥ 5 kg) at planning and end of the RT treatment

| Parameters | At RT planning | At the end of RT |
|---------------------------|----------------|-------------------|
| Area under the curve | 0.77 | 0.82 ^a |
| Sensitivity | 0.77 | 0.98 |
| Specificity | 0.67 | 0.59 |
| Positive predictive value | 0.43 | 0.46 |
| Negative predictive value | 0.90 | 0.99 |

Abbreviation as in Table 1.

^a Statistical significance at RT planning and at the end of RT.

In traditional statistical modeling, we have an assumption of a data distribution first. The models with estimated parameters are validated particularly by goodness-of-fit tests. Machine learning, based on traditional statistics, allows more variables to be taken into account, and its algorithms suggest what is most important without those assumptions. Model validation is measured by its prediction accuracy particularly by cross-validation. Because of the complex nature in the causal relationship of weight loss, we pursued the latter machine learning approach (CART model) in this paper and evaluated its feasibility in knowledge discovery.

We used the CART algorithm as a knowledge–discovery tool because of its interpretability. The CART model provides critical variables threshold and their directional influence on the outcomes, for example, reducing the dose to the superior constrictor muscle will lead to no critical weight loss (Fig 3). In fact, it catalyzed the sharing of ideas between data scientists and clinicians to bridge the understanding across the multidisciplinary research group. CART can be used to explore datasets and can readily identify interactions among prognostic factors at each node in the tree. Unlike logistic regression approaches, CART does not require a specification of the function that is used to model covariates. These strengths are especially valuable in addressing the data heterogeneity commonly associated with clinical datasets. Collectively, our results provide a rich basis for future hypotheses regarding weight loss prevention in the irradiated HNC patient.

CART is particularly sensitive to missing data, however, which can remain a challenge even for prospectively collected datasets such as ours. Additional algorithms are available that can address the influence of missing data on the accuracy of a machine learning prediction.⁴¹ Despite this, the model performances generated in our feasibility study demonstrated AUC >0.75, which suggests promising accuracy warranting further investigations. Another challenge is model unstableness and overfitting when applying CART. To balance the recursive partitioning (branching) and pruning to overcome the potential instability in model development, we used leave-one-out cross-validation to address this issue. Although some other algorithms are known to have potentially high performance in prediction, we chose CART algorithm because of its interpretability.

Random forest is a tree-based algorithm that is known to be less prone to overfitting. Random forest grows many classification trees (eg, CART) by using randomly selected bootstrap samples and randomly selected predictor variable sets.⁴² Each tree gives a classification (vote) for a sample, and the forest chooses the classification having the most votes over all the trees in the forest. The steps of random predictor selection and voting give this algorithm high robustness in prediction, but the interpretability of the selected predictors is reduced to understanding which predictors carry high importance; the algorithm does not

indicate the threshold and direction as the CART model does, making rationalization more difficult in the knowledge–discovery approach. Neural network is also subject to interpretability challenges⁴³; these methods are worthy of further exploration because their potential to improve prediction accuracy may outweigh their interpretability challenges.

Another possible concern is that the threshold cutoff for critical weight loss remains debated.⁴⁴ Several studies have used a 5% or 10%^{18,20,21,45,46} threshold as a definition of critical weight loss based on the international consensus statement of the Academy of Nutrition and Dietetics and American Society for Parenteral and Enteral Nutrition.⁴⁷ Our informatics platform with its data mining tools can readily be adapted to any threshold cutoff or definition of critical weight loss. In fact, we have built different CART weight loss prediction models using 5% as cutoff as an example (eFigs 1 and 2; available as supplementary material online only at www.practical.radonc.org, Table 1). Although they are not exactly the same tree models, we found common predictors (eFigs 1 and 2) that show robust associations with different weight-loss cutoffs that warrant attention in future big data approaches regarding weight loss prediction.⁴⁸ In our analysis, we considered the absolute loss (≥ 5 kg) as the primary cutoff and used the relative loss ($\geq 5\%$) as an auxiliary analysis. The commonly identified nodes are of particular interest because they may be important to weight loss regardless of cutoff. In summary, our flexible informatic infrastructure can readily facilitate building prediction models at multiple time points and cutoffs; more significantly, this lends itself to the iterative reanalysis of the models, especially by testing newly acquired factors that may be hypothesized to be relevant over time (Figs 1 and 2). This increases the probability of generating a valid, generalizable, and insightful model.

Last, our findings suggest that the informatics infrastructure that we have piloted can generate a large and diverse clinical dataset that can form the basis for advanced predictive modeling. We observed model 2 demonstrated that, with additional treatment toxicity information in the irradiated HNC patient, the accuracy of the prediction increased, suggesting that this approach has validity. The incorporation of more patient-specific information toward the end of RT improved the predictive accuracy of the weight loss model (AUC comparison between at planning and EOT: 5-kg model, $P = .03$; 5% model, $P = .08$) and is an encouraging indication that this infrastructure may facilitate the construction of a learning health system for precision radiation medicine. Our results were cross-validated by using a dataset in our institution. Demonstrating external validity as with traditional regression modeling approaches requires a comparable dataset to be available to then test the models that are generated. Further validation with other institutions' data needs to be investigated.

Conclusion

We have demonstrated the feasibility of creating 2 weight loss predictive models at different time points with improving receiver operating characteristics with incremental data. This reflects the potential of machine learning models in knowledge–discovery and potentially decision support from prospectively collected data in routine clinical workflow. Combined with large-scale analytic approaches, we believe that the Oncospace informatics framework can provide the foundation for developing a personalized learning health system, especially with the development of real-time clinical decision support systems.

Supplementary data

Supplementary material for this article (<https://doi.org/10.1016/j.adro.2017.11.006>) can be found at www.practicalradonc.org.

References

- Benedict SH, El Naqa I, Klein EE. Introduction to big data in radiation oncology: Exploring opportunities for research, quality assessment, and clinical care. *Int J Radiat Oncol*. 2016;95:871-872.
- Potters L, Ford E, Evans S, Pawlicki T, Mutic S. A systems approach using big data to improve safety and quality in radiation oncology. *Int J Radiat Oncol Biol Phys*. 2016;95:885-889.
- Jagsi R, Bekelman JE, Chen A, et al. Considerations for observational research using large data sets in radiation oncology. *Int J Radiat Oncol Biol Phys*. 2014;90:11-24.
- Mayo CS, Deasy JO, Chera BS, Freymann J, Kirby JS, Hardenberg PH. How can we effect culture change toward data-driven medicine? *Int J Radiat Oncol Biol Phys*. 2016;95:916-921.
- Huser V, Cimino JJ. Impending challenges for the use of big data. *Int J Radiat Oncol Biol Phys*. 2016;95:890-894.
- McNutt TR, Moore KL, Quon H. Needs and challenges for big data in radiation oncology. *Int J Radiat Oncol* 2016;95:909-915.
- Sloan JA, Halyard M, El Naqa I, Mayo C. Lessons from large-scale collection of patient-reported outcomes: Implications for big data aggregation and analytics. *Int J Radiat Oncol Biol Phys*. 2016;95:922-929.
- Benedict SH, Hoffman K, Martel MK, et al. Overview of the American Society for Radiation Oncology–National Institutes of Health–American Association of Physicists in Medicine Workshop 2015: Exploring opportunities for radiation oncology in the era of big data. *Int J Radiat Oncol Biol Phys*. 2016;95:873-879.
- Bowers MR, McNutt TR. Oncospace Consortium: A shared radiation oncology database system designed for personalized medicine and research. *Int J Radiat Oncol* 2015;93:E385.
- Robertson SP, Quon H, Kiess AP, et al. A data-mining framework for large scale analysis of dose–outcome relationships in a database of irradiated head and neck cancer patients. *Med Phys* 2015;42:4329–4337.
- Breiman L. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Stat Sci*. 2001;16:199-231.
- Lambin P, Roelofs E, Reymen B, et al. “Rapid Learning health care in oncology” – An approach towards decision support systems enabling customised radiotherapy. *Radiother Oncol*. 2013;109:159-164.
- Lambin P, Zindler J, Vanneste BGL, et al. Decision support systems for personalized and participative radiation oncology. *Adv Drug Deliv Rev*. 2017;109:131-153.
- Lustberg T, van Soest J, Jochems A, et al. Big data in radiation therapy: Challenges and opportunities. *Br J Radiol*. 2016;90:20160689.
- Vlooswijk CP, van Rooij PHE, Kruize JC, Schuring HA, Al-Mamgani A, de Roos NM. Dietary counselling and nutritional support in oropharyngeal cancer patients treated with radiotherapy: Persistent weight loss during 1-year follow-ups. *Eur J Clin Nutr*. 2016;70:54-59.
- Langius JAE, Zandbergen MC, Eerenstein SEJ, et al. Effect of nutritional interventions on nutritional status, quality of life and mortality in patients with head and neck cancer receiving (chemo)radiotherapy: A systematic review. *Clin Nutr*. 2013;32:671-678.
- Chasen MR, Bhargava R. A descriptive review of the factors contributing to nutritional compromise in patients with head and neck cancer. *Support Care Cancer*. 2009;17:1345-1351.
- Langius JAE, Bakker S, Rietveld DHF, et al. Critical weight loss is a major prognostic indicator for disease-specific survival in patients with head and neck cancer receiving radiotherapy. *Br J Cancer*. 2013;109:1093-1099.
- Langius JAE, van Dijk AM, Doornaert P, et al. More than 10% weight loss in head and neck cancer patients during radiotherapy is independently associated with deterioration in quality of life. *Nutr Cancer*. 2013;65:76-83.
- Mallick I, Gupta SK, Ray R, et al. Predictors of weight loss during conformal radiotherapy for head and neck cancers—How important are planning target volumes? *Clin Oncol*. 2013;25:557-563.
- Langius JAE, Twisk J, Kampman M, et al. Prediction model to predict critical weight loss in patients with head and neck cancer during (chemo)radiotherapy. *Oral Oncol*. 2016;52:91-96.
- Dueck AC, Mendoza TR, Mitchell SA, et al. Validity and reliability of the US National Cancer Institute’s patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *JAMA Oncol*. 2015;1:1051-1059.
- Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy scale: Development and validation of the general measure. *J Clin Oncol*. 1993;11:570-579.
- Chen AY, Frankowski R, Bishop-Leone J, et al. The development and validation of a dysphagia-specific quality-of-life questionnaire for patients with head and neck cancer: The M. D. Anderson dysphagia inventory. *Arch Otolaryngol Head Neck Surg*. 2001;127:870-876.
- Basch E, Reeve BB, Mitchell SA, et al. Development of the National Cancer Institute’s Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). *J Natl Cancer Inst*. 2014;106:dju244.
- Wallace KL, Middleton S, Cook IJ. Development and validation of a self-report symptom inventory to assess the severity of oral-pharyngeal dysphagia. *Gastroenterology*. 2000;118:678-687.
- Petit SF, Wu B, Kazhdan M, et al. Increased organ sparing using shape-based treatment plan optimization for intensity modulated radiation therapy of pancreatic adenocarcinoma. *Radiother Oncol*. 2012;102:38-44.
- Kazhdan M, Simari P, McNutt T, et al. A shape relationship descriptor for radiation therapy planning. *Med Image Comput Comput Assist Interv*. 2009;12:100-108.
- Wu B, Ricchetti F, Sanguineti G, et al. Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Med Phys*. 2009;36:5497-5505.
- Wu B, Ricchetti F, Sanguineti G, et al. Data-driven approach to generating achievable dose–volume histogram objectives in intensity-modulated radiotherapy planning. *Int J Radiat Oncol*. 2011;79:1241-1247.

31. Breiman L, Friedman J, Stone C. *Classification and Regression Trees (Wadsworth Statistics/Probability)*. New York: Chapman and Hall/CRC; 1984.
32. Hapfelmeier A, Hothorn T, Ulm K. Recursive partitioning on incomplete data using surrogate decisions and multiple imputation. *Comput Stat Data Anal*. 2012;56:1552-1565.
33. Feelders A. Handling missing data in trees: Surrogate splits or statistical imputation? In: Żytkow JM, Rauch J, eds. *Principles of Data Mining and Knowledge Discovery*, Vol. 1704. Berlin, Heidelberg: Springer Berlin Heidelberg; 1999:329-334.
34. Therneau T, Atkinson B, Brian R. Rpart: Recursive partitioning and regression trees. 2015. Available at: <https://cran.r-project.org/web/packages/rpart/index.html>. Accessed January 2, 2016.
35. Jager-Wittenaar H, Dijkstra PU, Vissink A, Laan BFAM, van der Oort RP, van Roodenburg JLN. Critical weight loss in head and neck cancer—prevalence and risk factors at diagnosis: An explorative study. *Support Care Cancer*. 2007;15:1045-1050.
36. Righini C-A, Timi N, Junet P, Bertolo A, Reyt E, Atallah I. Assessment of nutritional status at the time of diagnosis in patients treated for head and neck cancer. *Eur Ann Otorhinolaryngol Head Neck Dis*. 2013;130:8-14.
37. Farhangfar A, Makarewicz M, Ghosh S, et al. Nutrition impact symptoms in a population cohort of head and neck cancer patients: Multivariate regression analysis of symptoms on oral intake, weight loss and survival. *Oral Oncol*. 2014;50:877-883.
38. Wang X, Eisbruch A. IMRT for head and neck cancer: Reducing xerostomia and dysphagia. *J Radiat Res (Tokyo)*. 2016;57:i69-i75.
39. Hamlet S, Faull J, Klein B, et al. Mastication and swallowing in patients with postirradiation xerostomia. *Int J Radiat Oncol Biol Phys*. 1997;37:789-796.
40. Nguyen NP, Moltz CC, Frank C, et al. Dysphagia following chemoradiation for locally advanced head and neck cancer. *Ann Oncol*. 2004;15:383-388.
41. Batista GEAPA, Monard MC. An analysis of four missing data treatment methods for supervised learning. *Applied Artificial Intelligence*. 2003;17:519-533.
42. Breiman L. Random forests. *Mach Learn*. 2001;45:5-32.
43. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol*. 1996;49:1225-1231.
44. Dechaphunkul T, Martin L, Alberda C, Olson K, Baracos V, Gramlich L. Malnutrition assessment in patients with cancers of the head and neck: A call to action and consensus. *Crit Rev Oncol Hematol*. 2013;88:459-476.
45. Nourissat A, Bairati I, Samson E, et al. Predictors of weight loss during radiotherapy in patients with stage I or II head and neck cancer. *Cancer*. 2010;116:2275-2283.
46. Palazzi M, Tomatis S, Orlandi E, et al. Effects of treatment intensification on acute local toxicity during radiotherapy for head and neck cancer: Prospective observational study validating CTCAE, version 3.0, scoring system. *Int J Radiat Oncol Biol Phys*. 2008;70:330-337.
47. White JV, Guenter P, Jensen G, et al. Consensus statement: Academy of Nutrition and Dietetics and American Society for Parenteral and Enteral Nutrition characteristics recommended for the identification and documentation of adult malnutrition (undernutrition). *JPEN J Parenter Enteral Nutr*. 2012;36:275-283.
48. Rosenstein BS, Capala J, Efstathiou JA, et al. How will big data improve clinical and basic research in radiation therapy? *Int J Radiat Oncol Biol Phys*. 2016;95:895-904.