



OPEN ACCESS

EDITED BY

Xinyuan Chen,
Chinese Academy of Medical Sciences and
Peking Union Medical College, China

REVIEWED BY

Hong Qi Tan,
National Cancer Centre Singapore, Singapore
Michael Andrew Samuels,
Banner Health, United States

*CORRESPONDENCE

Satomi Shiraishi

✉ shiraishi.satomi@mayo.edu

RECEIVED 29 November 2023

ACCEPTED 14 February 2024

PUBLISHED 28 February 2024

CITATION

Madhavan S, Gamez M, Garces YI, Lester SC,
Ma DJ, Mundy DW, Neben Wittich MA,
Qian J, Routman DM, Foote RL and Shiraishi S
(2024) Impact of radiation dose distribution
on nutritional supplementation needs in
head and neck cancer radiotherapy: a
voxel-based machine learning approach.
Front. Oncol. 14:1346797.
doi: 10.3389/fonc.2024.1346797

COPYRIGHT

© 2024 Madhavan, Gamez, Garces, Lester, Ma,
Mundy, Neben Wittich, Qian, Routman, Foote
and Shiraishi. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Impact of radiation dose distribution on nutritional supplementation needs in head and neck cancer radiotherapy: a voxel-based machine learning approach

Sudharsan Madhavan, Mauricio Gamez, Yolanda I. Garces, Scott C. Lester, Daniel J. Ma, Daniel W. Mundy, Michelle A. Neben Wittich, Jing Qian, David M. Routman, Robert L. Foote and Satomi Shiraishi*

Department of Radiation Oncology, Mayo Clinic, Rochester, MN, United States

Objectives: To investigate the relationship between nutritional supplementation and radiation dose to the pharyngeal constrictor muscles and larynx for head and neck (HN) cancer patients undergoing radiotherapy.

Methods: We retrospectively analyzed radiotherapy (RT) dose for 231 HN cancer patients, focusing on the pharyngeal constrictors and larynx. We defined nutritional supplementation as feeding tube utilization or >10% weight loss from baseline within 90 days after radiotherapy completion. Using deformable image registration (DIR), we mapped each patient's anatomical structures to a reference coordinate system, and corresponding deformations were applied to dose matrices. Voxel doses were utilized as features for ridge logistic regression models, optimized through 5-fold cross-validation. Model performance was assessed with area under the curve of a receiver operating curve (AUC) and F1 score. We built and compared models using 1) pharyngeal constrictor voxels, 2) larynx voxels, 3) clinical factors and mean regional dose metrics, and 4) clinical factors and dose-volume histogram metrics. Test set AUCs were compared among the models, and feature importance was evaluated.

Results: DIR of the pharyngeal constrictors and larynx yielded mean Dice coefficients of 0.80 and 0.84, respectively. Pharyngeal constrictors voxels and larynx voxel models had AUC of 0.88 and 0.82, respectively. Voxel-based dose modeling identified the superior to middle regions of the pharyngeal constrictors and the superior region of larynx as most predictive of feeding tube use/weight loss. Univariate analysis found treatment setting, treatment laterality, chemotherapy, baseline dysphagia, weight, and socioeconomic status predictive of outcome. An aggregated model using mean doses of pharyngeal constrictors and larynx subregions had an AUC of 0.87 and the model using conventional DVH metrics had an AUC of 0.85 with p-value of 0.04. Feature importance calculations from the regional dose model indicated that mean doses to the superior-middle

pharyngeal constrictor muscles followed by mean dose to the superior larynx were most predictive of nutritional supplementation.

Conclusions: Machine learning modeling of voxel-level doses enables identification of subregions within organs that correlate with toxicity. For HN radiotherapy, doses to the superior-middle pharyngeal constrictors are most predictive of feeding tube use/weight loss followed by the doses to superior portion of the larynx.

KEYWORDS

voxel-based analysis, head and neck cancer, outcomes modeling, feeding tube, explainable machine learning, larynx, pharyngeal constrictor muscles, weight loss

1 Introduction

Swallowing difficulties are a prevalent side effect of radiotherapy (RT) treatments for head and neck cancers (1–5). RT involves targeting cancers with a three-dimensional (3D) radiation dose. This often leads to the incidental irradiation of nearby organs that play a role in swallowing. In contemporary practices, radiotherapy treatment plans condense the 3D dose distributions inside delineated organs into two-dimensional dose-volume histograms (DVHs). Specific metrics for organs at risk (OARs) and target volumes are scrutinized to reduce the chance of adverse side effects. Physicians qualitatively assess the spatial distribution of radiation doses, focusing on regions that may pose potential toxicities or affect target coverage. This method has proven effective, with a number of toxicities linked to DVH metrics through various normal tissue complication probability (NTCP) modeling, as evidenced in QUANTEC and various clinical trials (1, 5–10). For example, Mavroidis et al. (11), showed that generalized mean dose for superior pharyngeal constrictors to be most predictive of dysphagia at 6 months post-RT. Using Lyman Kutcher Burman (LKB) model which is a popular methodology accounting for seriality of the OAR (12, 13), their study reported D_{50} (dose at which there is 50% chance of complication) of 62.0 Gy, slope parameter $m = 0.1$, and dose-volume parameter $n = 0.49$ with AUC of 0.74 for superior pharyngeal constrictor muscles, suggesting moderate sensitivity to subregion damage (14). However, when relying exclusively on DVHs for analysis, delineation of the OARs is necessary, and it is presupposed that every section of an organ has an equal sensitivity to radiation and contributes identically to the overall risk of toxicity. Additionally, Samant et al. reported that machine learning (ML) models can often quantify NTCP better than LKB models, motivating explorations of ML approaches for toxicity analysis (15).

Furthermore, recent studies have shown that the toxicity observed can depend on which subregion of a segmented organ was irradiated (2–4, 16–37). Functionally distinct subregions within a single OAR contour may not be accounted for in current

treatment planning. For example, Jiang et al. used ML to identify subregions within the parotid and submandibular glands that correlated with xerostomia (22). Another study by Eisbruch et al. found the pharyngeal constrictor muscles and glottic/supraglottic larynx subsites were most dysphagia-related using videofluoroscopy (38). In clinical practice, it is common for glottic, supraglottic, and subglottic larynx subregions to be grouped under a single larynx segmentation. Likewise, the superior, middle, and inferior pharyngeal constrictor muscles are collectively evaluated as one entity during treatment planning. In addition to the potential presence of distinct subregions, there is a growing number of studies on voxel-based optimization of radiation treatment plans (39–41). These studies have used voxel-based objectives for optimization and opens an opportunity to reflect spatial dose constraints during treatment planning.

Conventionally, investigation of sub-regions of an OAR required radiomics texture extraction based on images and/or time-consuming and resource-intensive manual contouring of each region (42). Manual contouring is also subject to inter-observer variations and inconsistent implementation of contouring guidelines (43–45). In this study, we employed a machine-learning approach to determine the radiation dose's voxel-wise correlation with malnutrition, pinpointing distinctive features in the pharyngeal constrictors and larynx sub-regions without preliminary contouring. We focused on interpretable feature importance to analyze the spatial dose dependency within an organ. This may eventually inform dose sparing of a subregion during treatment planning to reduce malnutrition risk. While methodologies akin to ours have been used in head and neck RT studies focusing on toxicities such as xerostomia and acute dysphagia (28, 29), to the best of our knowledge, there has not been an investigation into the 3D dose distribution's impact on malnutrition, measured by feeding tube (FT) utilization and weight loss. Furthermore, predicting the need for a feeding tube early in radiotherapy can also enhance patient care. Although prophylactic FT placement is occasionally advised (46), its blanket application can be unnecessary or harmful. About half of the preemptively

inserted FTs prove marginally beneficial (47), with complications such as infections being prevalent (48). Moreover, FT placement postpones the transition back to regular diets (49), adversely affecting long-term well-being (18, 50, 51). Therefore, identifying patients in true need of FT can improve quality of care.

2 Materials and methods

2.1 Cohort selection

After Institutional Review Board approval, we conducted a retrospective analysis of 352 patients treated for head and neck cancer at our institution from January 2016 to November 2020, all of whom had granted consent for their medical records to be used in research. We accessed our department's patient outcomes database (52) and filtered for patients based on International Classification of Diseases (ICD)-9 and 10 codes (53, 54) specific to cancers in the salivary glands, oropharynx, oral cavity, nasopharynx, nasal cavity, sinuses, larynx, and hypopharynx. We excluded patients based on the following criteria: absence of baseline dysphagia assessment before radiotherapy (60 patients), FT insertion before RT (34 patients), radiation doses outside the range of 1.2-2.2 Gy/fraction, discernible disfigurement of the pharyngeal constrictors and larynx due to disease or surgery, and prior RT in an area with potential overlap (27 patients). Ultimately, 231 patients, treated with either photon or proton radiotherapy with prescription dose in the range of 30-81.6 Gy delivered in 15-68 fractions, were deemed suitable for this study.

2.2 Data collection

Table 1 summarizes the clinical variables studied in our analysis. We sourced data from our institution's electronic health record reporting database, focusing on parameters such as gender, feeding tube usage, weight, birth date, and the primary address's 9-digit zip code. Feeding tube utilization data was gathered by looking for procedure codes corresponding to the insertion of stomach, gastrostomy, or jejunostomy tubes. Patients who had a feeding tube before starting RT were excluded based on the procedure date, as such utilization is likely attributable to surgery, disease, or both as opposed to RT. Socioeconomic status was inferred using the Area Deprivation Index by Kind and Buckingham (55), derived from the zip code associated with the patient's primary address. This index, ranging from 1 to 100, gauges the socioeconomic disadvantage of a neighborhood, with higher scores denoting greater disadvantage. We also collected baseline dysphagia grades assessed by the care team following the Common Terminology Criteria for Adverse Events (CTCAE v4.03) within ± 2 weeks of radiotherapy initiation (56, 57). Additionally, data on smoking habits, concurrent chemotherapy, and treatment context (either primary or post-operative RT) was obtained via chart reviews. Table 1 provides information on clinical variables and on treatment sites of the patients included in the study.

All radiotherapy treatments were planned in the Eclipse treatment planning system (Siemens Healthineers company,

TABLE 1 List of clinical parameters investigated, and treatment site of the patient cohort considered in this study.

Clinical variables		Value
Baseline dysphagia grade	Grades 0-1	208
	Grades 2-3	23
Age at RT start	Mean	60.85 years
	Range	[23, 89] years
ADI	Mean	45.01
	Range	[1, 98]
Treatment setting	Surgery before RT	146
	No surgery before RT	85
Smoker status	Smoker	120
	Never smoker	111
Gender	Male	189
	Female	42
Radiation type	Photon	131
	Proton	100
Concurrent chemotherapy	Yes	156
	No	75
Treatment site (% of FT/WL)	Salivary glands	6 (50%)
	Oropharynx	155 (54%)
	Oral cavity	23 (61%)
	Nasopharynx	12 (75%)
	Nasal cavity and sinuses	15 (33%)
	Larynx	17 (35%)
	Hypopharynx	3 (67%)

Erlangen, Germany). Physical radiation doses in proton plans were scaled by 1.1 (58-60) to account for relative biological effect compared to photon plans. Photon treatments were generally planned to use two to four volumetric modulated arcs. Proton treatments were planned with a pencil beam scanning method utilizing two to five static fields. From our planning system, we exported DICOM files corresponding to radiation doses, CT scans, and structural sets. Of note, the head and neck anatomy of these data sets were retrospectively segmented consistently by a specially trained team of physicians and medical dosimetry assistant as part of a separate project (61). The contouring followed the consensus guidelines in Brouwer et al. (62).

2.3 Endpoint definition

The endpoint used to characterize malnutrition was feeding tube (FT) utilization or >10% weight loss from baseline within 90 days after radiotherapy completion. Though FT usage typically arises when a patient sheds over 10% of their initial weight, some

clinicians and/or patients opt against it on a case-by-case basis. Our practice is to only recommend a FT if it is clinically indicated (typically >10% weight loss from baseline). We do not use FT prophylactically. Regardless, such patients remain malnourished, indicating a toxicity affecting their quality of life. We chose to merge these observations and study them as one endpoint; for the remainder of the paper, this endpoint will be referred to as FT/WL for simplicity. The baseline weight was defined as the weight recorded closest to the RT commencement, ensuring it was within a ±2-week window of that date.

2.4 Analysis

2.4.1 Overview

As illustrated in Figure 1, this analysis had two different methodologies after data processing. In the first methodology, 3D voxel-based dose models were trained to identify regions within the pharyngeal constrictors and larynx that better differentiated toxicity endpoints. Because the voxel-based model utilized a large number of input voxels (22,020 for pharyngeal constrictors and 20,814 for larynx), it was prone to overfitting. To validate and confirm the subregion findings, the second methodology investigated models with reduced features where mean doses from segmented subregions of the pharyngeal constrictor muscles and larynx along with clinical variables (14 input features). Similarly, another model that combined DVH metrics used in our clinic with the identical clinical parameters (13 input features) was studied as a comparison.

2.4.2 Data preprocessing

All patients' pharyngeal constrictors and larynx were deformably registered to a reference patient's corresponding organs to align them in the same coordinate system. The reference patient shown in Figure 2, was chosen based on having a larynx structure size close to the population average, as well as the absence of any disfigurement due to prior surgery or disease involving the pharyngeal constrictors and/or larynx. Deformable registration was performed using the open-source package, Elastix (63–65). The planning images of most patients included in the study had a native resolution of 1.27 x 1.27 mm² in the axial direction, with slice thicknesses of either 1 or 2 mm. To maintain uniformity, all images were interpolated to a resolution of 1.27 mm x 1.27 mm x 1 mm in the coronal, sagittal, and axial directions, respectively. The DICOM images (66) were then cropped around organ +4 mm using the open-source packages DicomRTTool and ANTs (67, 68). Rigid, affine, and deformable transformations were subsequently applied. The quality of image registrations was evaluated using the Dice coefficient. For the 3D models, we also augmented the training dataset by flipping the OAR contours and dose left-right to create mirrored dose distributions. This is based on the premise that both the pharyngeal constrictors and larynx are midline structures and there are no laterality preferences for one versus the other. The flipped contour and doses were registered to the reference coordinate system in the same manner as the original data. After the deformable image registrations, the same deformation was applied to each patient's dose matrix to obtain dose within the reference coordinate system. Dose to each voxel was extracted, and

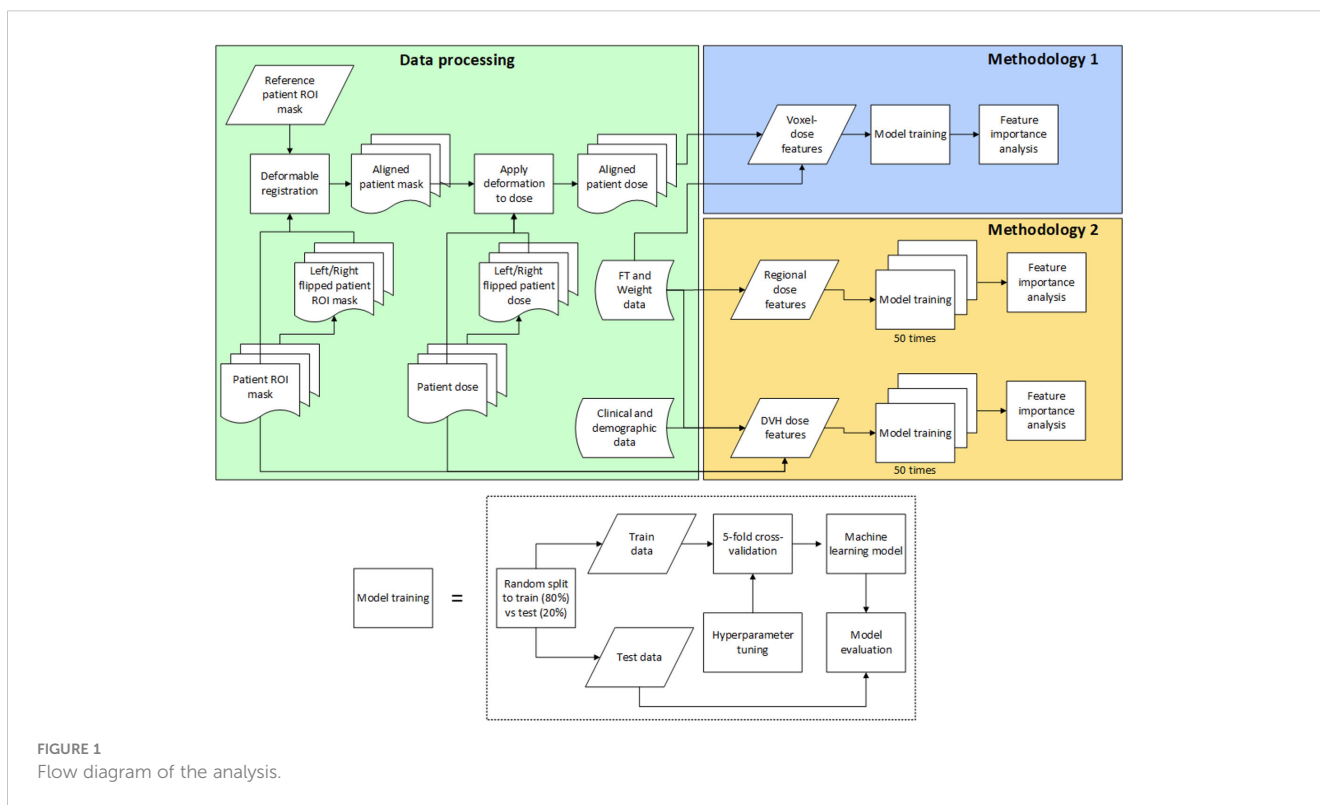


FIGURE 1
Flow diagram of the analysis.



each numerical input feature was standardized. This standardization process involved adjusting each feature such that it had a mean of zero and a standard deviation of one among the training set. Standardization ensures that our model's performance is not biased by variations in the absolute dose levels, but rather focuses on the relative differences in dose distribution. All categorical features were one-hot encoded and expressed in terms of zeros and ones.

2.4.3 Modeling and statistical analysis

We initially evaluated ridge logistic regression, eXtreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM) algorithms for modeling. Ridge logistic regression was implemented using the open-source package *cuML* (69) and *scikit-learn* (70), while XGBoost (71) and LightGBM (72) models were utilized from their respective repositories. We performed voxel-based analysis using these algorithms with a subset of cohort, and areas under the curves (AUCs) for receiver operating curves (ROCs) were compared. The feature importance maps from XGBoost and lightGBM were sparse and their performance were comparable, so we opted to utilize ridge logistic regression for the remainder of the study.

In Methodology 1, voxel doses were the only input features for the models, with the pharyngeal constrictors and larynx analyzed in separate models. The data set was randomly split into training and test sets in an 80-20% ratio, and the same training patients were used for both pharyngeal constrictors and larynx models. The hyperparameter was tuned using the *Optuna* package (73) over the 5-fold cross-validation, with ROC AUC as the performance metric. Accuracy and F1 scores evaluated at the threshold of 0.5 were also utilized to compare the general performance of the models. Full training data was used to perform the final fit with the optimized hyperparameter to create the model. For the 3D models, the standardized input features allowed the coefficients

from the ridge logistic regression to directly signify feature importance. Specifically, each coefficient from the regression showcases the change in log odds of the outcome for a unit increase in its respective feature. These feature importance coefficients were qualitatively reviewed in a scatter plot for any sub-regions that differentiated FT/WL better than others. In addition to ridge logistic regression, other regularized methods such as LASSO and elastic net were also attempted. However, these L1-based regularization, promoting sparsity combined with highly correlated input parameters yielded spatially fragmented and variable feature importance maps. Therefore, ridge regression was preferred to retain spatial coherence for subregion identification.

In methodology 2, a univariate logistic regression was first performed for each clinical and demographic variable fitting to the binary outcomes of FT/WL, and p-value and odds ratio were calculated. We first developed a model exclusively incorporating the non-dosimetric clinical variables listed in Table 2, which served to establish a baseline for prediction performance. Variables identified as predictive of FT/WL in the univariate analysis were used for modeling with the dosimetric variables. In the final models, there were seven to eight dosimetric variables and six clinical and demographic variables. The clinical and demographic variables were common between the DVH model and the regional dose model: area deprivation index, baseline weight, treatment setting (primary/post-operative), concurrent chemotherapy (yes/no), bilateral treatment (yes/no), and baseline dysphagia grade 0 (yes/no). The dosimetric variables for the DVH metrics model were mean larynx dose, mean pharyngeal constrictor dose, larynx V50Gy and V60Gy, and pharyngeal constrictor V50Gy, V55Gy, and V60Gy, where $VxGy$ represents the percent volume of the organ covered by x Gy or more. These DVH metrics were chosen because they are used to evaluate head and neck treatment plans in our clinic. For the regional dose model, the pharyngeal constrictors were divided into superior, middle and inferior pharyngeal constrictors on the reference patient as well as midline and lateral regions as shown in Figure 3C, grouping the regions based on the feature importance from the voxel models. Following a similar grouping, the larynx was divided into supraglottic larynx and inferior regions as shown in Figure 3F. Mean doses from these subregions were used as the dosimetric variables for the regional dose model. Using these input features, models were trained 50 times with the 5-fold cross-validation process with varying random data splits to confirm that the learned hyperparameters and feature importances were not heavily dependent on a specific random partitioning of the dataset. To assess feature importance, we employed a permutation test (74). We randomized the data of one input variable at a time and evaluated the resulting drop in model performance. The magnitude of performance decline, as measured by the change in ROC AUC from the unaltered data, indicates the importance of that variable to the model's predictive capability. We conducted this permutation process 50 times for every variable, incorporating 5-fold cross-validation on the training dataset in each iteration. The average change in the AUC was used to assess the difference in performance, and the two-tailed Wilcoxon rank sum test was used to evaluate the statistical significance of this change. To compare the general performance

TABLE 2 P-values and odds ratios from the univariate analysis of non-dosimetric variables correlating with FT/WL.

Non-dosimetric variables	P-value	Odds Ratios [95% CI]
Treatment setting (primary vs post-operative treatment)	<0.0001	4.62 [2.55 – 8.38]
Bilateral treatment	<0.0001	6.72 [3.07 – 14.73]
Concurrent chemotherapy	0.001	2.54 [1.44 – 4.49]
Baseline weight	0.01	1.014 [1.000-1.028]
Area deprivation index	0.02	1.014 [1.002-1.026]
Baseline dysphagia grade = 0	0.03	0.51 [0.28 – 0.92]
Gender	0.16	1.63 [0.83-3.20]
Treatment modality (photon vs proton)	0.17	0.69 [0.41 – 1.17]
Never smoker	0.34	0.78 [0.46 – 1.30]
Age	0.39	0.99 [0.97-1.01]

P-values and odds ratios from the univariate analysis of nondosimetric variables correlating with FT/WL. Variables with p-values below 0.05 (bold) showed a statistically significant correlation with FT/WL placement.

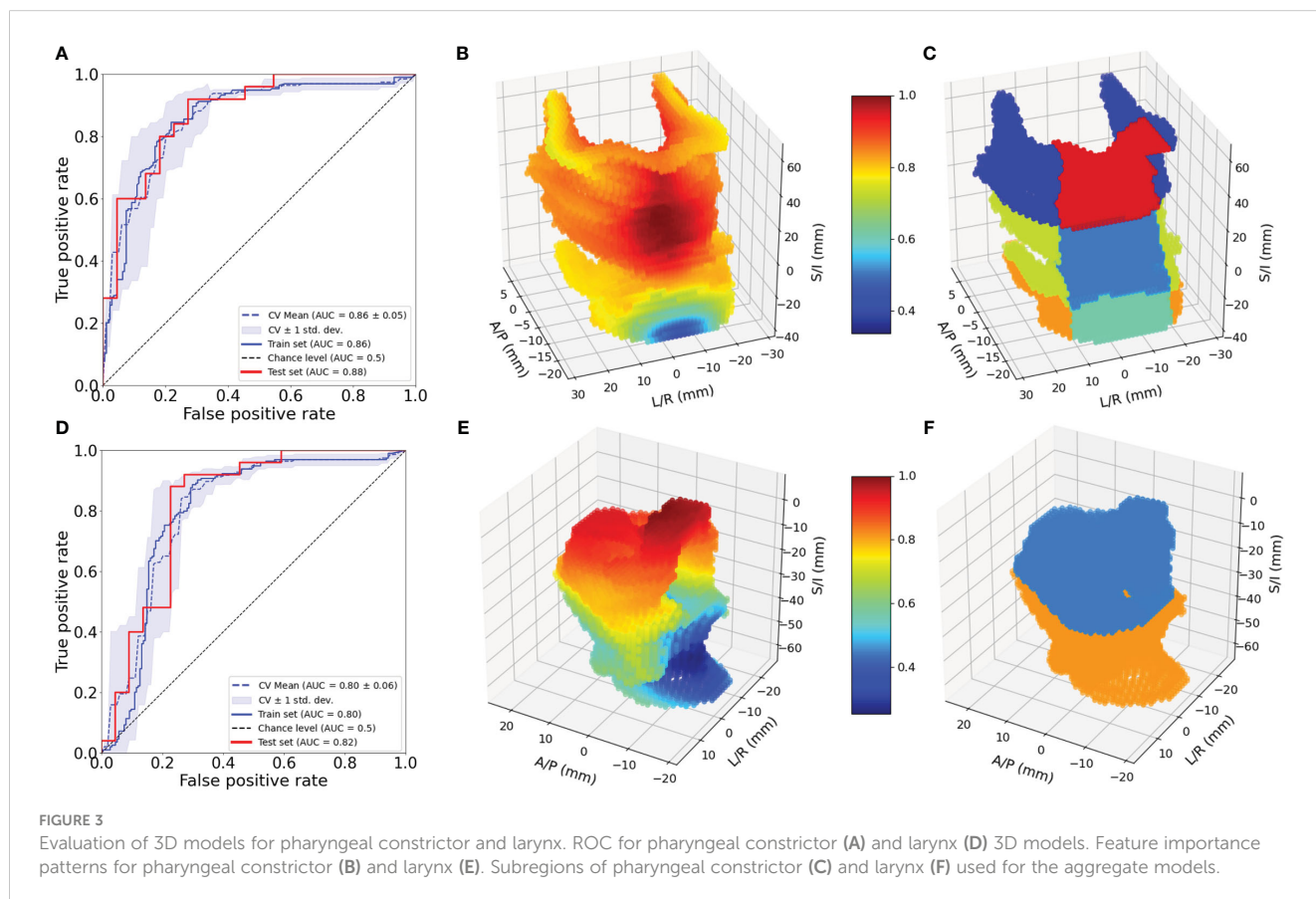
of the regional dose model and the DVH metrics model, test set AUCs from the 50 trials were also compared using the two-tailed Wilcoxon rank sum test.

3 Results

Of the 231 patients, 64 patients were found to have utilized FT, and 106 patients lost more than 10% of their baseline weight within 90 days after completing RT. Combining these observations, 122 patients had either FT, WL or both. All registrations in this study yielded a Dice coefficient of greater than 0.7, with a mean and standard deviation of 0.80 ± 0.02 for pharyngeal constrictors and 0.84 ± 0.06 for larynx. Figure 4A, B shows the mean and standard deviation of pharyngeal constrictor dose distributions for the entire cohort. The average mean pharyngeal constrictor dose was 52.3 ± 11.3 Gy and 31.7 ± 15.2 Gy for those who experienced FT/WL and those who did not, respectively. Figure 4C, D show the mean and the standard deviation of larynx dose distributions. Among those who experienced FT/WL, the average mean larynx dose was 41.6 ± 14.4 Gy while that of patients who did not experience FT/WL was 26.5 ± 18.3 Gy.

3.1 Voxel-based dose model

Figure 3A shows the ROC curves for the 3D pharyngeal constrictor model. The AUC for the models' training and test data were found to be 0.87 and 0.86, respectively, with cross-validation yielding 0.86 ± 0.02 . The accuracy and F1 scores for the test data were 0.83 and 0.85, respectively. Figure 3B shows the feature importance from the model, highlighting the middle to superior pharyngeal constrictors as being more impactful when



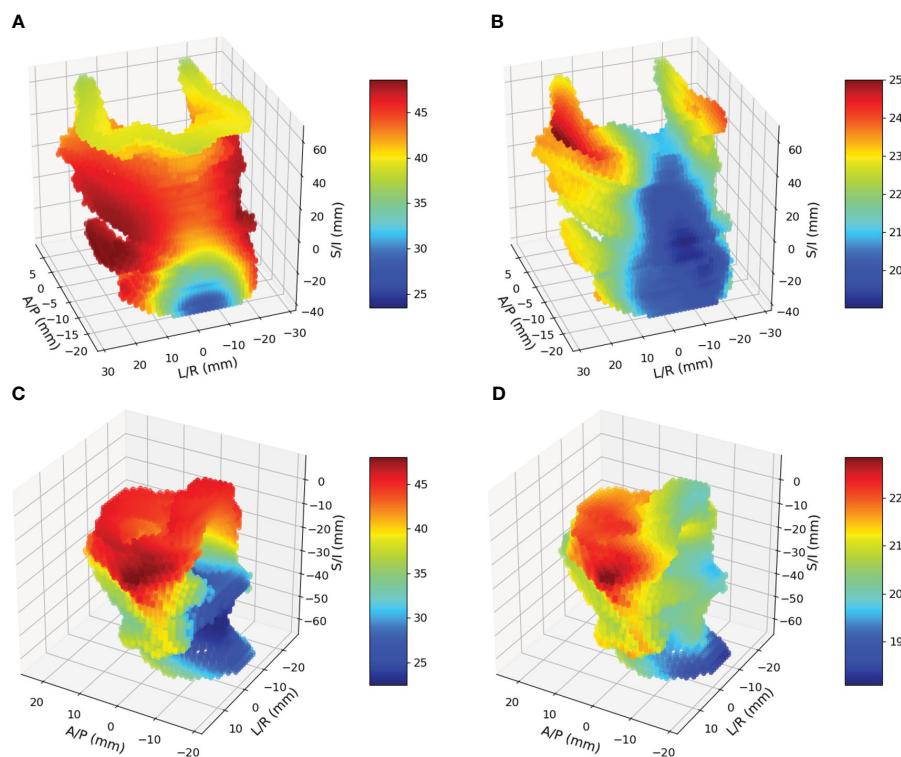


FIGURE 4
(A) and **(B)** are mean and standard deviation of pharyngeal constrictor doses. **(C)** and **(D)** are mean and standard deviations of larynx doses.

compared to the inferior region. The midline voxels of the superior and middle pharyngeal constrictors were most important for determining FT/WL for this cohort. **Figure 3D, E** illustrates ROC and the feature importance pattern from the larynx 3D model. The AUC for the model’s training and test data was 0.80 and 0.82, respectively, with cross-validation yielding 0.80 ± 0.06 . The accuracy and F1 scores for the test data were both 0.72. The feature importance consistently highlights the superior part of the larynx as a critical factor for predicting FT/WL.

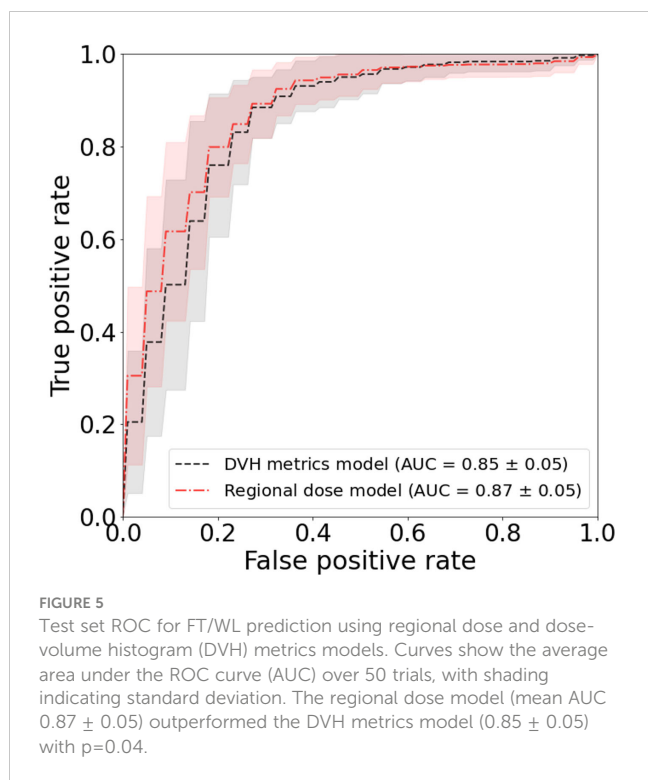
3.2 Univariate analysis and aggregated dosimetric model

The results of the univariate fits to non-dosimetric variables are shown in **Table 2**. There were six variables that showed significant correlation with FT/WL: treatment setting (primary/post-operative), treatment laterality (bilateral/unilateral), concurrent chemotherapy (yes/no), baseline dysphagia grade =0, baseline weight, and ADI. These six non-dosimetric variables were used as part of the aggregated dosimetric models. Gender, treatment modality (photon vs proton), smoking status and age did not show significant correlation with FT/WL. A clinical variables-only model using the non-dosimetric variables in **Table 2** achieved an AUC of 0.76 ± 0.01 and 0.74 ± 0.06 for training and test set, respectively. The pharyngeal constrictor and larynx subregions are illustrated in **Figure 3C, F**, respectively. Test set ROCs from the

regional dose and DVH metrics models are shown in **Figure 5**. The lines indicate the average of the 50 trials, and the shaded area indicates the standard deviations. The regional dose model, with a mean AUC of 0.87 ± 0.05 , demonstrated marginally superior performance compared to the DVH metrics model, which had a mean AUC of 0.85 ± 0.05 with a p-value of 0.04. The F1 scores for both models were 0.82 ± 0.04 . The incorporation of dosimetric features in the DVH and regional dose models significantly improved performance over the clinical variables-only model ($p < 0.0001$). **Table 3** shows the results of a permutation test to evaluate feature importance. Input features that significantly affected the performance of the model when shuffled—as assessed by AUC compared to the actual data—are listed in order of significance. For the regional dose model, the most important features were pharyngeal constrictor doses, particularly the superior to middle region. While dose to the superior part of the larynx significantly contributed to the performance of the regional dose model, none of the larynx DVH metrics showed significant performance gain in the DVH metrics model.

4 Discussion

This study delves into the relationship between radiation dose distributions to two organs at risk (OARs) and feeding tube use or weight loss in head and neck cancer patients receiving radiotherapy. Employing deformable image registration and ridge logistic



regression, we mapped structures and identified key regions, particularly the superior part of the pharyngeal constrictors and the superior part of the larynx, as crucial determinants for FT/WL. This finding was revealed with 3D voxel-based models and validated using aggregated dose model with reduced input features to minimize risk of overfitting.

Our study highlights that dose to the superior pharyngeal constrictor muscles was most important in predicting FT/WL in our cohort. While we did not find studies investigating the same endpoint, our findings are in general agreement with prior clinical (75) and outcomes studies (43) demonstrating the radiosensitivity of sub-regions in organs and their role in post-radiation dysphagia

and aspiration. Feng et al. (76) and Eisbruch et al. (38) also reported that highest correlations of videofluoroscopy based aspiration and dysphagia to the superior pharyngeal constrictor in a prospective study. Petras et al. (77) have evaluated the relationship between dose to larynx subregions and swallowing toxicities assessed by aspiration at one year, and Hedstrom et al. (78) considered dysphagia at 6-months post-treatment. Both studies identified the epiglottis as a critical subregion, in line with our findings.

Overall, the performance of the DVH metrics and the regional dose models were similar. This supports our current standard of care of using DVH for treatment planning. However, we note that feature importance analysis for the DVH metrics model revealed that the model performance did not significantly depend on larynx DVH metrics. On the other hand, the aggregated regional dose model utilized mean doses from superior-middle pharyngeal constrictor and supraglottic larynx. This study motivates further investigation into dose sparing of these subregions. Understanding the dose distribution effects will facilitate voxel-based optimization, evaluation and interpretation of treatment plans that have similar DVH metrics.

The strength of our work includes the number of cohorts, standardized contouring of the pharyngeal constrictors and the larynx, as well as inclusion of patients treated with proton therapy. In contrast to the pharyngeal constrictor and larynx segmentation created during clinical workflow, contours used in this study were retrospectively drawn to achieve high consistency and conformance to the contouring guideline. Consequently, the definition of the organ was consistent across the entire patient cohort. For future studies, automatic segmentation using artificial intelligence-driven algorithms along with quality assurance processes could aid in generating more consistent anatomical segmentations than those available from clinical data. However, automatic segmentation algorithms are typically not trained to segment substructures. As this study demonstrated, a voxel-based approach allows us to eliminate the need for exhaustive contouring of each substructure *a priori* for the entire cohort, thereby streamlining the analytical process.

TABLE 3 Evaluation of feature importance based on a permutation test for the DVH metrics and regional dose models. The table lists input features in descending order of significance, based on their impact on model performance assessed by ROC AUC. (PC: Pharyngeal constrictor muscles).

	Features from DVH metrics model	Mean AUC loss	Features from regional dose model	Mean AUC loss
1	PC: Mean dose	-0.021	PC: superior lateral mean dose	-0.022
2	PC: V50Gy (%)	-0.016	PC: superior midline mean dose	-0.020
3	PC: V55Gy (%)	-0.012	PC: middle midline mean dose	-0.011
4	Baseline weight	-0.006	PC: middle lateral mean dose	-0.009
5	PC: V65Gy (%)	-0.006	Baseline weight	-0.004
6	Area Deprivation Index	-0.004	Larynx: superior mean dose	-0.004
7	Concurrent chemotherapy (yes/no)	-0.002	Area Deprivation Index	-0.003
8	Treatment setting	-0.001	Concurrent chemotherapy (yes/no)	-0.002
9	Baseline dysphagia grade = 0 (yes/no)	-0.0006	Baseline dysphagia grade = 0 (yes/no)	-0.0004

Limitations of our work include the nature of single-institution studies and the lack of a treatment planning component to the study, so that it remains to be seen what amount of dose reduction to the mid to superior pharyngeal constrictor and supraglottic larynx could be achieved without compromising treatment quality. An additional limitation is the lack of controlled study on treatment modality (proton/photon). The relative biological effectiveness (RBE) of 1.1 (58–60) for protons used in this study applies to tumor control outcomes, which may not directly translate to functional outcomes in normal tissues. Furthermore, our study included 21 photon patients who received lower prescription doses of 30 Gy. Since our primary objective was to investigate effects of various dose distributions, we opted to include those treatments to increase diversity. However, these choices likely confounded the analysis with respect to treatment modality. While promising, our findings warrant validation in diverse cohorts, treatment modality, treatment planning techniques, and treatment regimens. Nevertheless, our study highlights the potential of interpretable voxel-based modeling to elucidate impact of inhomogeneous dose distributions within an organ.

5 Conclusion

In conclusion, the 3D voxel-based analysis and the aggregated regional dose analysis highlighted the superior subregion of the pharyngeal constrictor muscles and the supraglottic larynx as the most important predictor of FT/WL within 90 days of RT.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: protected health information. Requests to access these datasets should be directed to shiraishi.satomi@mayo.edu.

Ethics statement

The studies involving humans were approved by Mayo Clinic Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

References

1. Caglar HB, Tishler RB, Othus M, Burke E, Li Y, Goguen L, et al. Dose to larynx predicts for swallowing complications after intensity-modulated radiotherapy. *Int J Radiat Oncol Biol Phys.* (2008) 72:1110–8. doi: 10.1016/j.ijrobp.2008.02.048
2. Fua TF, Corry J, Milner AD, Cramb J, Walsham SF, Peters LJ. Intensity-modulated radiotherapy for nasopharyngeal carcinoma: clinical correlation of dose to the pharyngo-esophageal axis and dysphagia. *Int J Radiat Oncol Biol Phys.* (2007) 67:976–81. doi: 10.1016/j.ijrobp.2006.10.028
3. Schwartz DL, Hutcheson K, Barringer D, Tucker SL, Kies M, Holsinger FC, et al. Candidate dosimetric predictors of long-term swallowing dysfunction after

Author contributions

SM: Data curation, Investigation, Methodology, Visualization, Writing – original draft. MG: Writing – review & editing. YG: Writing – review & editing. SL: Writing – review & editing. DJM: Writing – review & editing. DWM: Writing – review & editing. MW: Writing – review & editing. JQ: Methodology, Writing – review & editing. DR: Writing – review & editing. RF: Conceptualization, Data curation, Writing – review & editing. SS: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was primarily funded by Varian Medical Systems and partly funded by Mayo Clinic Eagle's Pilot Award.

Acknowledgments

This is a short text to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

oropharyngeal intensity-modulated radiotherapy. *Int J Radiat Oncol Biol Phys.* (2010) 78:1356–65. doi: 10.1016/j.ijrobp.2009.10.002

4. Humbert-Vidan L, Patel V, Andlauer R, King AP, Guerrero Urbano T. Prediction of mandibular ORN incidence from 3D radiation dose distribution maps using deep learning. *Lecture Notes Comput Sci (including subseries Lecture Notes Artif Intell Lecture Notes Bioinformatics).* (2022) 13540 LNCS:49–58. doi: 10.1007/978-3-031-17721-7_6/COVER

5. Inada M, Nishimura Y, Ishikura S, Ishikawa K, Murakami N, Kodaira T, et al. Organs-at-risk dose constraints in head and neck intensity-modulated radiation

- therapy using a dataset from a multi-institutional clinical trial (JCOG1015A1). *Radiat Oncol.* (2022) 17:133. doi: 10.1186/s13014-022-02105-3
6. Bentzen SM, Constine LS, Deasy JO, Eisbruch A, Jackson A, Marks LB, et al. Quantitative analyses of normal tissue effects in the clinic (QUANTEC): an introduction to the scientific issues. *Int J Radiat OncologyBiologyPhys.* (2010) 76:S3–9. doi: 10.1016/j.ijrobp.2009.09.040
7. Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of normal tissue complication probability models in the clinic. *Int J Radiat Oncol Biol Phys.* (2010) 76:S10–19. doi: 10.1016/j.ijrobp.2009.07.1754
8. Brodin NP, Kabarriti R, Garg MK, Guha C, Tomé WA. Systematic review of normal tissue complication models relevant to standard fractionation radiation therapy of the head and neck region published after the QUANTEC reports. *Int J Radiat OncologyBiologyPhys.* (2018) 100:391–407. doi: 10.1016/j.ijrobp.2017.09.041
9. 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. doi: 10.1016/j.radonc.2011.08.009
10. Rancati T, Schwarz M, Allen AM, Feng F, Popovtzer A, Mittal B, et al. Radiation dose–volume effects in the larynx and pharynx. *Int J Radiat Oncol Biol Phys.* (2010) 76:S64–9. doi: 10.1016/j.ijrobp.2009.03.079
11. Mavroidis P, Price A, Fried D, Kostich M, Amdur R, Mendenhall W, et al. Dose–volume toxicity modeling for de-intensified chemo-radiation therapy for HPV-positive oropharynx cancer. *Radiother Oncol.* (2017) 124:240–7. doi: 10.1016/j.radonc.2017.06.020
12. Kutcher GJ, Burman C, Brewster L, Goitein M, Mohan R. Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations. *Int J Radiat Oncol Biol Phys.* (1991) 21:137–46. doi: 10.1016/0360-3016(91)90173-2
13. Lyman JT. Complication probability as assessed from dose–volume histograms. *Radiat Res Suppl.* (1985) 8:S13–19. doi: 10.2307/3583506
14. Seppenwoolde Y, Lebesque JV, de Jaeger K, Belderbos JSA, Boersma LJ, Schilstra C, et al. Comparing different NTCP models that predict the incidence of radiation pneumonitis. Normal tissue complication probability. *Int J Radiat Oncol Biol Phys.* (2003) 55:724–35. doi: 10.1016/S0360-3016(02)03986-X
15. Samant P, de Ruyscher D, Hoebbers F, Canters R, Hall E, Nutting C, et al. Machine learning for normal tissue complication probability prediction: Predictive power with versatility and easy implementation. *Clin Trans Radiat Oncol.* (2023) 39:100595. doi: 10.1016/j.ctro.2023.100595
16. Buettner F, Miah AB, Gulliford SL, Hall E, Harrington KJ, Webb S, et al. Novel approaches to improve the therapeutic index of head and neck radiotherapy: an analysis of data from the PARSPORT randomised phase III trial. *Radiother Oncol.* (2012) 103:82–7. doi: 10.1016/j.radonc.2012.02.006
17. Cella L, Monti S, Thor M, Rimner A, Deasy JO, Palma G. Radiation-induced dyspnea in lung cancer patients treated with stereotactic body radiation therapy. *Cancers.* (2021) 13:3734. doi: 10.3390/cancers13153734
18. Dohopolski M, Wang K, Morgan H, Sher D, Wang J. Use of deep learning to predict the need for aggressive nutritional supplementation during head and neck radiotherapy. *Radiother Oncol.* (2022) 171:129–38. doi: 10.1016/j.radonc.2022.04.016
19. Ebert MA, Gulliford S, Acosta O, de Crevoisier R, McNutt T, Heemsbergen WD, et al. Spatial descriptions of radiotherapy dose: normal tissue complication models and statistical associations. *Phys Med Biol.* (2021) 66:12TR01. doi: 10.1088/1361-6560/ac0681
20. Guo Y, Jiang W, Lakshminarayanan P, Han P, Cheng Z, Bowers M, et al. Spatial radiation dose influence on xerostomia recovery and its comparison to acute incidence in patients with head and neck cancer. *Adv Radiat Oncol.* (2020) 5:221–30. doi: 10.1016/j.adro.2019.08.009
21. Isaksson LJ, Pepa M, Zaffaroni M, Marvaso G, Alterio D, Volpe S, et al. Machine learning-based models for prediction of toxicity outcomes in radiotherapy. *Front Oncol.* (2020) 10:790. doi: 10.3389/fonc.2020.00790
22. Jiang W, Lakshminarayanan P, Hui X, Han P, Cheng Z, Bowers M, et al. Machine learning methods uncover radiomorphologic dose patterns in salivary glands that predict xerostomia in patients with head and neck cancer. *Adv Radiat Oncol.* (2019) 4:401–12. doi: 10.1016/j.adro.2018.11.008
23. Kobayashi K, Murakami N, Takahashi K, Inaba K, Igaki H, Hamamoto R, et al. A population-based statistical model for investigating heterogeneous intraprostatic sensitivity to radiation toxicity after 125I Seed Implantation. *In Vivo.* (2019) 33:2103–11. doi: 10.21873/invivo.11710
24. Murr M, Brock KK, Fusella M, Hardcastle N, Hussein M, Jameson MG, et al. Applicability and usage of dose mapping/accumulation in radiotherapy. *Radiother Oncol.* (2023) 182:109527. doi: 10.1016/j.radonc.2023.109527
25. Palma G, Monti S, Conson M, Pacelli R, Cella L. Normal tissue complication probability (NTCP) models for modern radiation therapy. *Semin Oncol.* (2019) 46:210–8. doi: 10.1053/j.seminoncol.2019.07.006
26. Palma G, Monti S, Cella L. Voxel-based analysis in radiation oncology: A methodological cookbook. *Physica Med.* (2020) 69:192–204. doi: 10.1016/j.ejmp.2019.12.013
27. Voshart DC, Wiedemann J, van Luijk P, Barazzuol L. Regional responses in radiation-induced normal tissue damage. *Cancers.* (2021) 13:1–21. doi: 10.3390/CANCERS13030367
28. Bourbonne V, Lucia F, Jaouen V, Bert J, Rehn M, Pradier O, et al. Development and prospective validation of a spatial dose pattern based model predicting acute pulmonary toxicity in patients treated with volumetric arc-therapy for locally advanced lung cancer. *Radiother Oncol.* (2021) 164:43–9. doi: 10.1016/j.radonc.2021.09.008
29. Bourbonne V, Lucia F, Jaouen V, Pradier O, Visvikis D, Schick U. Combination of radiomics features and functional radiosensitivity enhances prediction of acute pulmonary toxicity in a prospective validation cohort of patients with a locally advanced lung cancer treated with VMAT-radiotherapy. *J Personal Med.* (2022) 12:1926. doi: 10.3390/jpm12111926
30. Cella L, Monti S, Xu T, Luzzi R, Stanzione A, Durante M, et al. Probing thoracic dose patterns associated to pericardial effusion and mortality in patients treated with photons and protons for locally advanced non-small-cell lung cancer. *Radiother Oncol.* (2021) 160:148–58. doi: 10.1016/j.radonc.2021.04.025
31. Heemsbergen WD, Incrocci L, Pos FJ, Heijmen BJM, Witte MG. Local dose effects for late gastrointestinal toxicity after hypofractionated and conventionally fractionated modern radiotherapy for prostate cancer in the HYPRO trial. *Front Oncol.* (2020) 10:469. doi: 10.3389/fonc.2020.00469
32. Monti S, Xu T, Liao Z, Mohan R, Cella L, Palma G. On the interplay between dosimics and genomics in radiation-induced lymphopenia of lung cancer patients. *Radiother Oncol.* (2022) 167:219–25. doi: 10.1016/j.radonc.2021.12.038
33. Monti S, Xu T, Mohan R, Liao Z, Palma G, Cella L. Radiation-induced esophagitis in non-small-cell lung cancer patients: voxel-based analysis and NTCP modeling. *Cancers.* (2022) 14:1833. doi: 10.3390/cancers14071833
34. Palma G, Monti S, Pacelli R, Liao Z, Deasy JO, Mohan R, et al. Radiation pneumonitis in thoracic cancer patients: multi-center voxel-based analysis. *Cancers.* (2021) 13:3553. doi: 10.3390/cancers13143553
35. Shelley LEA, Sutcliffe MPF, Thomas SJ, Noble DJ, Romanchikova M, Harrison K, et al. Associations between voxel-level accumulated dose and rectal toxicity in prostate radiotherapy. *Phys Imaging Radiat Oncol.* (2020) 14:87–94. doi: 10.1016/j.phro.2020.05.006
36. Veiga C, Lim P, Anaya VM, Chandy E, Ahmad R, D'Souza D, et al. Atlas construction and spatial normalisation to facilitate radiation-induced late effects research in childhood cancer. *Phys Med Biol.* (2021) 66:105005. doi: 10.1088/1361-6560/abf010
37. Wilson LJ, Bryce-Atkinson A, Green A, Li Y, Merchant TE, van Herk M, et al. Image-based data mining applies to data collected from children. *Physica Med.* (2022) 99:31–43. doi: 10.1016/j.ejmp.2022.05.003
38. Eisbruch A, Kim HM, Feng FY, Lyden TH, Haxer MJ, Feng M, et al. Chemo-IMRT of oropharyngeal cancer aiming to reduce dysphagia: swallowing organs late complication probabilities and dosimetric correlates. *Int J Radiat Oncol Biol Phys.* (2011) 81:e93–99. doi: 10.1016/j.ijrobp.2010.12.067
39. Mai Y, Kong F, Yang Y, Zhou L, Li Y, Song T. Voxel-based automatic multi-criteria optimization for intensity modulated radiation therapy. *Radiat Oncol.* (2018) 13:241. doi: 10.1186/s13014-018-1179-7
40. Cotrutz C, Xing L. Using voxel-dependent importance factors for interactive DVH-based dose optimization. *Phys Med Biol.* (2002) 47:1659–69. doi: 10.1088/0031-9155/47/10/304
41. Eisbruch A, Schwartz M, Rasch C, Vineberg K, Damen E, Van As CJ, et al. Dysphagia and aspiration after chemoradiotherapy for head-and-neck cancer: Which anatomic structures are affected and can they be spared by IMRT? *Int J Radiat OncologyBiologyPhys.* (2004) 60:1425–39. doi: 10.1016/j.ijrobp.2004.05.050
42. Desideri I, Loi M, Francolini G, Becherini C, Livi L, Bonomo P. Application of radiomics for the prediction of radiation-induced toxicity in the IMRT era: current state-of-the-art. *Front Oncol.* (2020) 10:1708. doi: 10.3389/fonc.2020.01708
43. Pepin MD, Anaya S, Foote RL, Garces Y, Hosfield EA, Mundy D, et al. Modeling patient-reported swallowing outcomes with dose to the epiglottis. *Int J Radiat OncologyBiologyPhys.* (2021) 111:e377. doi: 10.1016/j.ijrobp.2021.07.1108
44. Apolle R, Appold S, Bijl HP, Blanchard P, Bussink J, Faviere-Finn C, et al. Inter-observer variability in target delineation increases during adaptive treatment of head-and-neck and lung cancer. *Acta Oncol.* (2019) 58:1378–85. doi: 10.1080/0284186X.2019.1629017
45. Petkar I, McQuaid D, Dunlop A, Tyler J, Hall E, Nutting C. Inter-observer variation in delineating the pharyngeal constrictor muscle as organ at risk in radiotherapy for head and neck cancer. *Front Oncol.* (2021) 11:644767. doi: 10.3389/fonc.2021.644767
46. Rutter CE, Yovino S, Taylor R, Wolf J, Cullen KJ, Ord R, et al. Impact of early percutaneous endoscopic gastrostomy tube placement on nutritional status and hospitalization in patients with head and neck cancer receiving definitive chemoradiation therapy. *Head Neck.* (2011) 33:1441–7. doi: 10.1002/hed.21624
47. Madhoun MF, Blankenship MM, Blankenship DM, Krempel GA, Tierney WM. Prophylactic PEG placement in head and neck cancer: How many feeding tubes are unused (and unnecessary)? *World J Gastroenterol.* (2011) 17:1004–8. doi: 10.3748/wjg.v17.i8.1004
48. Grant DG, Bradley PT, Pothier DD, Bailey D, Caldera S, Baldwin DL, et al. Complications following gastrostomy tube insertion in patients with head and neck cancer: a prospective multi-institution study, systematic review and meta-analysis. *Clin Otolaryngol.* (2009) 34:103–12. doi: 10.1111/j.1749-4486.2009.01889.x
49. Langmore S, Krisciunas GP, Miloro KV, Evans SR, Cheng DM. Does PEG use cause dysphagia in head and neck cancer patients? *Dysphagia.* (2012) 27:251–9. doi: 10.1007/s00455-011-9360-2

50. Beaver ME, Matheny KE, Roberts DB, Myers JN. Predictors of weight loss during radiation therapy. *Otolaryngology-Head Neck Surg.* (2001) 125:645–8. doi: 10.1067/mhn.2001.120428
51. Cady J. Nutritional support during radiotherapy for head and neck cancer: the role of prophylactic feeding tube placement. *Clin J Oncol Nurs.* (2007) 11:875–80. doi: 10.1188/07.CJON.875-880
52. Whitaker TJ, Mayo CS, Ma DJ, Haddock MG, Miller RC, Corbin KS, et al. Data collection of patient outcomes: one institution's experience. *J Radiat Res.* (2018) 59:i19–24. doi: 10.1093/jrr/try013
53. Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. International Classification of Diseases, 9th Revision, Clinical Modification codes in discharge abstracts are poor measures of complication occurrence in medical inpatients. *Med Care.* (1997) 35(6):589–602. doi: 10.1097/00005650-199706000-00005
54. Hedegaard H, Johnson RL, Garnett M, Thomas KE. The International classification of diseases, 10th revision, clinical modification (ICD–10–CM): external cause-of-injury framework for categorizing mechanism and intent of injury. (2019) (136):1–22.
55. Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible — The neighborhood atlas. *N Engl J Med.* (2018) 378:2456–8. doi: 10.1056/NEJMp1802313
56. Institute NC. Common Terminology Criteria for Adverse Events (CTCAE) (2021). Available online at: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm.
57. Eisbruch A, Lyden T, Bradford CR, Dawson LA, Haxer MJ, Miller AE, et al. Objective assessment of swallowing dysfunction and aspiration after radiation concurrent with chemotherapy for head-and-neck cancer. *Int J Radiat OncologyBiologyPhys.* (2002) 53:23–8. doi: 10.1016/S0360-3016(02)02712-8
58. Paganetti H, Niemierko A, Ancukiewicz M, Gerweck LE, Goitein M, Loeffler JS, et al. Relative biological effectiveness (RBE) values for proton beam therapy. *Int J Radiat OncologyBiologyPhys.* (2002) 53:407–21. doi: 10.1016/S0360-3016(02)02754-2
59. Paganetti H. Relating proton treatments to photon treatments via the relative biological effectiveness—Should we revise current clinical practice? *Int J Radiat Oncol Biol Phys.* (2015) 91:892–4. doi: 10.1016/j.ijrobp.2014.11.021
60. Paganetti H, Blakely E, Carabe-Fernandez A, Carlson DJ, Das IJ, Dong L, et al. Report of the AAPM TG-256 on the relative biological effectiveness of proton beams in radiation therapy. *Med Phys.* (2019) 46:e53–78. doi: 10.1002/mp.13390
61. Tryggstad E, Anand A, Beltran C, Brooks J, Cimmiyotti J, Grimaldi N, et al. Scalable radiotherapy data curation infrastructure for deep-learning based autosegmentation of organs-at-risk: A case study in head and neck cancer. *Front Oncol.* (2022) 12:936134. doi: 10.3389/fonc.2022.936134
62. Brouwer CL, Steenbakkers RJHM, Bourhis J, Budach W, Grau C, Grégoire V, et al. CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines. *Radiother Oncol.* (2015) 117:83–90. doi: 10.1016/j.radonc.2015.07.041
63. Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging.* (2010) 29:196–205. doi: 10.1109/TMI.2009.2035616
64. Shamonin DP, Bron EE, Lelieveldt BPF, Smits M, Klein S, Staring M, et al. Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease. *Front neuroinform.* (2014) 7:50. doi: 10.3389/fninf.2013.00050
65. Marstal K, Berendsen F, Staring M, Klein S. (2016). SimpleElastix: A user-friendly, multi-lingual library for medical image registration, *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 574–82.
66. Anderson BM, Wahid KA, Brock KK. Simple python module for conversions between DICOM images and radiation therapy structures, masks, and prediction arrays. *Pract Radiat Oncol.* (2021) 11:226–9. doi: 10.1016/j.prro.2021.02.003
67. Tustison NJ, Cook PA, Holbrook AJ, Johnson HJ, Muschelli J, Devenyi GA, et al. The ANTsX ecosystem for quantitative biological and medical imaging. *Sci Rep.* (2021) 11:9068. doi: 10.1038/s41598-021-87564-6
68. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage.* (2011) 54:2033–44. doi: 10.1016/j.neuroimage.2010.09.025
69. Raschka S, Patterson J, Nolet C. Machine learning in python: main developments and technology trends in data science, machine learning, and artificial intelligence. *Information.* (2020) 11:193. doi: 10.3390/info11040193
70. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in {Python}. *J Mach Learn Res.* (2011) 12:2825–30.
71. Chen T, Guestrin C. (2016). XGBoost: A scalable tree boosting system, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*, New York, NY, USA: Association for Computing Machinery, pp. 785–94. doi: 10.1145/2939672.2939785
72. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: A highly efficient gradient boosting decision tree. *Adv Neural Inf Process Syst.* (2017) 30:3149–57.
73. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. (2019). Optuna: A next-generation hyperparameter optimization framework, in: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage AK USA, pp. 2623–31. ACM. doi: 10.1145/3292500.3330701
74. Breiman L. Random forests. *Mach Learn.* (2001) 45:5–32. doi: 10.1023/A:1010933404324
75. Feng FY, Kim HM, Lyden TH, Haxer MJ, Worden FP, Feng M, et al. Intensity-modulated chemoradiotherapy aiming to reduce dysphagia in patients with oropharyngeal cancer: clinical and functional results. *J Clin Oncol.* (2010) 28:2732–8. doi: 10.1200/JCO.2009.24.6199
76. Feng FY, Kim HM, Lyden TH, Haxer MJ, Feng M, Worden FP, et al. Intensity-modulated radiotherapy of head and neck cancer aiming to reduce dysphagia: early dose-effect relationships for the swallowing structures. *Int J Radiat Oncol Biol Phys.* (2007) 68:1289–98. doi: 10.1016/j.ijrobp.2007.02.049
77. Petras KG, Rademaker AW, Refaat T, Choi M, Thomas TO, Pauloski BR, et al. Dose-volume relationship for laryngeal substructures and aspiration in patients with locally advanced head-and-neck cancer. *Radiat Oncol.* (2019) 14:49. doi: 10.1186/s13014-019-1247-7
78. Hedström J, Tuomi L, Finizia C, Olsson C. Identifying organs at risk for radiation-induced late dysphagia in head and neck cancer patients. *Clin Transl Radiat Oncol.* (2019) 19:87–95. doi: 10.1016/j.ctro.2019.08.005