



Original Article

Normal tissue complication probability model for acute oral mucositis in patients with head and neck cancer undergoing carbon ion radiation therapy based on dosimetry, radiomics, and dosiomics



Xiangdi Meng^a, Zhuojun Ju^a, Makoto Sakai^{b,*}, Yang Li^c, Atsushi Musha^{b,d}, Nobuteru Kubo^b, Hidemasa Kawamura^b, Tatsuya Ohno^{a,b}

^a Department of Radiation Oncology Gunma University Graduate School of Medicine Maebashi Japan

^b Gunma University Heavy Ion Medical Center Maebashi Japan

^c Department of Radiation Oncology Harbin Medical University Cancer Hospital Harbin China

^d Department of Oral and Maxillofacial Surgery and Plastic Surgery Gunma University Graduate School of Medicine Maebashi Japan

ARTICLE INFO

ABSTRACT

Keywords:

Carbon ion radiation therapy
Head and neck cancer
Prediction model
Radiomics
Dosiomics

Background and purpose: To develop a normal tissue complication probability (NTCP) model for predicting grade ≥ 2 acute oral mucositis (AOM) in head and neck cancer patients undergoing carbon-ion radiation therapy (CIRT).

Methods and materials: We retrospectively included 178 patients, collecting clinical, dose-volume histogram (DVH), radiomics, and dosiomics data. Patients were randomly divided into training (70%) and test sets (30%). Feature selection involved univariable logistic regression, least absolute shrinkage and selection operator regression, stepwise backward regression, and Spearman's correlation test, with the bootstrap method ensuring reliability. Multivariable models were built on the training set and evaluated using the test set.

Results: The optimal NTCP model incorporated a DVH parameter ($V_{37\text{Gy}}$ [relative biological effectiveness, RBE]), radiomics, and dosiomics features, achieving an area under the curve (AUC) of 0.932 in the training set and 0.959 in the test set. This hybrid model outperformed those based on single DVH, radiomics, dosiomics, or clinical data (Bonferroni-adjusted $p < 0.001$ and $\Delta\text{AUC} > 0$ for all comparisons in 1,000 bootstrap validations). Calibration curves showed strong agreement between predictions and outcomes. A 44.0 % AOM risk threshold was proposed, yielding accuracies of 87.1 % in the training set and 90.7 % in the test set.

Conclusions: We developed the first NTCP model for estimating AOM risk in head and neck cancer patients undergoing CIRT and proposed a risk stratification. This model may assist in clinical decision-making and improve treatment planning for AOM prevention and management by identifying high-risk patients.

Introduction

Acute oral mucositis (AOM) is a significant adverse event in the treatment of head and neck (HN) cancer, frequently resulting in pain, dysphagia, and treatment interruptions, with substantial impacts on treatment outcomes and patient quality of life (QOL) [1–5]. Despite dose-limiting strategies, AOM affects 80 to 100 % of HN cancer patients [6–9], with approximately 29 % experiencing severe mucositis [5]. Preserving the oral mucosa adjacent to the primary site remains technically challenging [5], necessitating early intervention and prediction.

Carbon ion radiotherapy (CIRT) is favored for its distinct Bragg peak

and high relative biological effectiveness (RBE), enabling targeted high doses to tumors while sparing normal tissues [10,11]. However, reports on CIRT-related oral mucositis are scarce, with incidence rates reported between 47 % and 65.2 % [3,12,13]. The absence of established dose-volume constraints and limited clinical evidence on CIRT-induced AOM highlight the need for effective preventive strategies in practice [14].

To our knowledge, no normal tissue complication probability (NTCP) model has been established for CIRT-induced AOM. Traditional models typically rely on parameters extracted from dose-volume histograms (DVHs) to predict AOM. However, these DVH parameters are

* Corresponding author at: Gunma University Heavy Ion Medical Center, 3-39-22, Showa-machi, Maebashi, Gunma 371-8511, Japan.

E-mail address: sakai-m@gunma-u.ac.jp (M. Sakai).

often highly correlated with one another, resulting in the final developed model being a univariable model with limited performance [15,16]. Dosiomics, which examines the spatial heterogeneity of three-dimensional (3D) dose distributions, provides deeper insights into irradiation toxicity, potentially enhancing NTCP models [17–21]. This approach has demonstrated superior predictive performance in various irradiation-induced toxicities compared with traditional models [17,18,22–24]. Moreover, radiomics, which analyzes subtle tissue differences, is crucial for understanding irradiation-induced injury in different oral mucosal structures [1,25,26]. Integrating radiomics and dosiomics allows for a more detailed analysis of mucosal dose distribution, potentially improving NTCP model accuracy [27,28].

The study aimed to develop the first individualized NTCP model for predicting CIRT-induced AOM (grade ≥ 2) to support clinical decision-making. Model building considers the clinical, DVH, radiomics, and dosiomics features to optimize AOM risk assessment. Additionally, the performance of this hybrid model was compared with that of traditional single models.

Methods and materials

Patients

This study, approved by the institutional review board (No. HS2023–202), retrospectively collected data from 194 consecutive patients with HN cancer who underwent CIRT between 2010 and 2020. Patients under 18 years, with a total oral mucosal dose < 0.1 Gy[RBE], intracranial tumors, unevaluable oral ulcers, recurrence at the same site, or irradiation fractions < 16 were excluded. A total of 178 patients, with a median age of 64 years (interquartile range: 52–74), were included and randomly divided into training (70 %) and test (30 %) sets (Table 1).

CIRT and toxicity assessment

The CIRT plan used a pencil-beam algorithm in the XiO-N system (Elekta Sweden, Mitsubishi Electric, Japan) (Figs. 1A–1C), supplemented by the K2DOSE calculation engine developed by the National Institute of Radiological Sciences (NIRS), Japan [3,16,29–31]. The clinical dose distribution was calculated using the HIMAC model [32,33]. This model, as previously reported, was based on experimental data from human salivary gland tumor cells to establish survival curves and to derive α and β parameters from linear energy transfer distributions calculated using the linear-quadratic model for RBE determination [13,29,31,34,35]. The clinical dose, expressed in Gy[RBE], was determined by multiplying the physical dose with the RBE value. The prescribed dose was 57.6, 64, or 70.4 Gy[RBE] in 16 fractions, depending on tumor type [3,16]. Patients were immobilized in the supine position using thermoplastic shells (Shellfitter, Kuraray, Osaka, Japan) in customized cradles (Moldcare, Alcare, Tokyo, Japan). A customized mouthpiece stabilized the mandible, facilitating the oral mucosa for evaluation and reducing side effects [3].

AOM was assessed by physicians or nurses before CIRT, every 4 fractions, at the end of treatment, and 90 days after treatment using the Common Terminology Criteria for Adverse Events (CTCAE, version 4.0). Here, we focused on grade ≥ 2 AOM, as these cases often require clinical interventions and are associated with delayed responses.

Oral mucosa surface contours (OMSC) segmentation

Oral mucosal thickness was defined as a 3-mm depth of the mucosal surface measured on computed tomography (CT) [36], with appropriate retraction near the bone (Figs. 1D–1F). This method ensured anatomical accuracy by excluding the muscular tissue [37–39]. The OMSC included the mucosa of the hard palate, lingual surface, lingual lateral margins, sublingual and lingual frenulum, cheeks, alveolar, gingiva, floor of the mouth, and internal lip. Additionally, parts of the pharyngeal mucosa

Table 1
Clinical characteristics of the patient (n = 178).

Characteristics	Overall n = 178	Training cohort n = 124	Test cohort n = 54	P value
AOM, n (%)				0.302
Grade 0–1	61 (34.3)	46 (37.1)	15 (27.8)	
Grade 2–3	117 (65.7)	78 (62.9)	39 (72.2)	
Age, years				
Median [IQR]	64.0 [52.0, 74.0]	65.5 [52.2, 74.0]	61.5 [52.0, 74.0]	0.389
≤ 70 years	113 (63.5)	81 (65.3)	32 (59.3)	0.546
> 70 years	65 (36.5)	43 (34.7)	22 (40.7)	
Gender, n (%)				0.718
Female	91 (51.1)	65 (52.4)	26 (48.1)	
Male	87 (48.9)	59 (47.6)	28 (51.9)	
Tumor location, n (%)				0.647
Nasal cavity	96 (53.9)	68 (54.8)	28 (51.9)	
Oral cavity	26 (14.6)	18 (14.5)	8 (14.8)	
Pharynx	15 (8.4)	9 (7.3)	6 (11.1)	
Parotid Gland and Ear	28 (15.7)	18 (14.5)	10 (18.5)	
Other ^a	13 (7.3)	11 (8.9)	2 (3.7)	
Stage, n (%)				0.474
I	7 (3.9)	6 (4.8)	1 (1.9)	
II	56 (31.5)	40 (32.3)	16 (29.6)	
III	21 (11.8)	12 (9.7)	9 (16.7)	
IV	94 (52.8)	66 (53.2)	28 (51.9)	
Chemotherapy, n (%)				0.830
No	145 (81.5)	100 (80.6)	45 (83.3)	
Yes	33 (18.5)	24 (19.4)	9 (16.7)	
Surgery, n (%)				1.000
No	139 (78.1)	97 (78.2)	42 (77.8)	
Yes	39 (21.9)	27 (21.8)	12 (22.2)	
Prescribed dose, [Gy (RBE)]				
57.6 (3.6 \times 16 fractions)	37 (20.8)	28 (22.6)	9 (16.7)	0.453
64.0 (4.0 \times 16 fractions)	124 (69.7)	86 (69.4)	38 (70.4)	
70.4 (4.4 \times 16 fractions)	17 (9.6)	10 (8.1)	7 (13.0)	
GTV, Median [IQR], (ml)	29.6 [13.0, 60.0]	31.4 [17.7, 51.8]	28.2 [12.4, 62.3]	0.662
Tumor Volume, Median [IQR], (ml)	31.5 [27.3, 35.8]	32.5 [27.4, 36.3]	31.2 [27.3, 35.1]	0.411

^a: Other includes maxillary, mandibular, osteosarcoma, liposarcoma, eye socket tumors.

Abbreviations: AOM, acute oral mucositis; IQR, interquartile range; GTV, gross tumor volume; RBE, relative biological effectiveness.

(lower surface of the soft palate and base of the tongue) were included for clinical observation.

Study design and statistical analysis

The study workflow is illustrated in Fig. 2. All variables included in the analysis were complete and did not require imputation or handling of missing data. DVH parameters were extracted using MIM Maestro software (version 7.3.2; MIM Software Inc., USA). Radiomics and dosiomics features were extracted from the DICOM files using the “PyRadiomics” library in Python (Version 3.7.1). Clinical features were obtained from medical records. The NTCP model was established using the following four steps.

Step 1. Clinical, dose-volume histogram, radiomics, and dosiomics feature extraction

The study included seven clinical variables: age, sex, tumor location, stage, chemotherapy, surgery, and fractionated dose. The oral mucosa volume exposed to CIRT was calculated in 1 Gy[RBE] increments. A total of 70 DVH parameters ($V_{1 \text{ Gy[RBE]}}$ to $V_{70 \text{ Gy[RBE]}}$) were selected for analysis.

Planning CT images were used for radiomic feature extraction. The CT scan parameters were as follows: tube voltage, 120 kVp; tube current, 250 mA; image size, 512 \times 512; pixel size, 0.877 mm; and slice thickness, 2 mm. All CT images, radiotherapy structures, and radiotherapy doses were resampled to an isotropic grid of 1 \times 1 \times 1 mm³ using

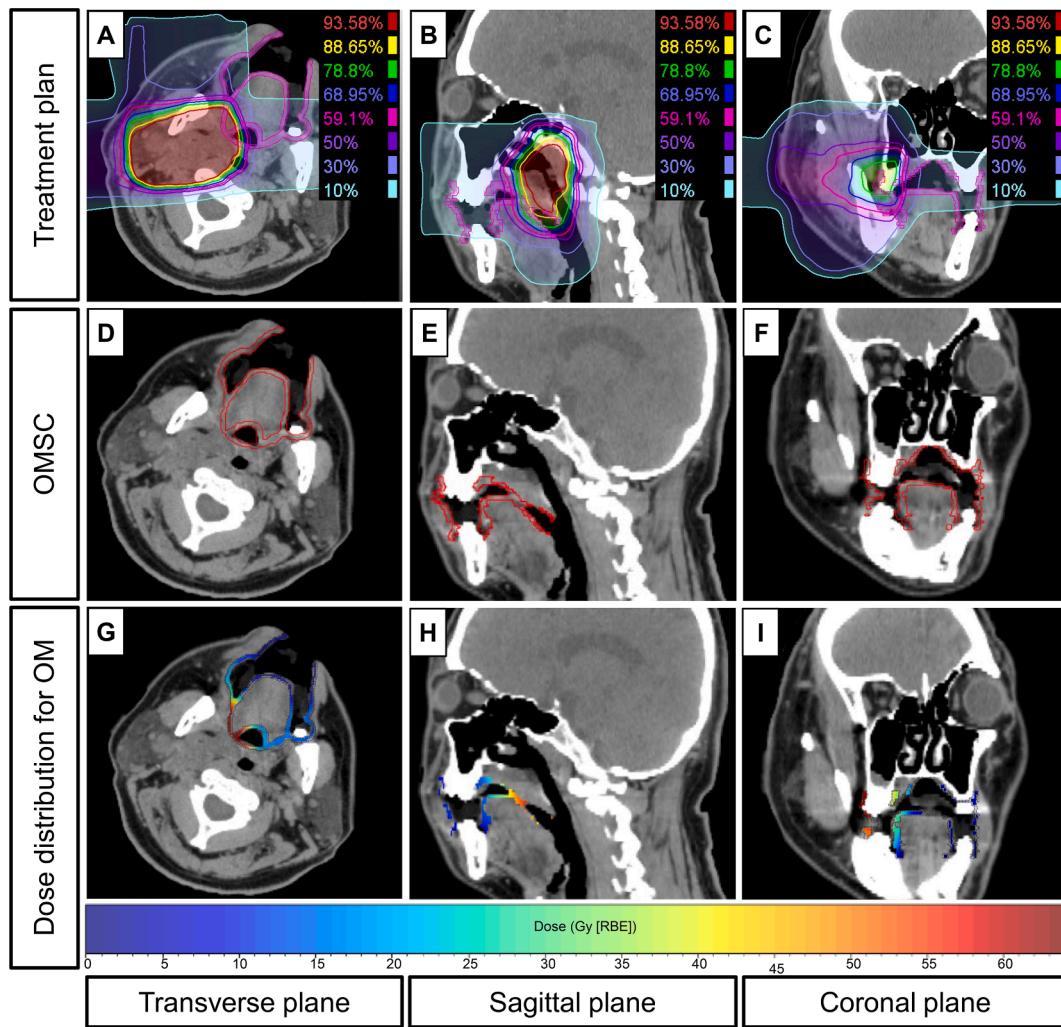


Fig. 1. Carbon ion radiotherapy planning (A-C), oral mucosal surface contours (OMCS) (D-F), and dose distribution to the oral mucosa (G-I) on transverse, coronal, and sagittal plane computed tomography images. (A) Carbon ion radiotherapy planning - Transverse plane. (B) Carbon ion radiotherapy planning - Coronal plane. (C) Carbon ion radiotherapy planning - Sagittal plane. (D) Oral mucosal surface contours (OMCS) - Transverse plane. (E) Oral mucosal surface contours (OMCS) - Coronal plane. (F) Oral mucosal surface contours (OMCS) - Sagittal plane. (G) Dose distribution to oral mucosa - Transverse plane. (H) Dose distribution to oral mucosa - Coronal plane. (I) Dose distribution to oral mucosa - Sagittal plane.

nearest-neighbor interpolation [20]. A total of 107 radiomic features were extracted from OMSC, including 14 shape, 18 first-order, and 75 texture features. The texture features were categorized as follows: 24 grey-level co-occurrence matrix (GLCM), 16 grey-level run length matrix (GLRLM), 16 grey-level size zone matrix (GLSZM), 14 grey-level dependence matrix (GLDM), and five features of the neighboring grey-tone difference matrix (NGTDM). Supplementary Table S1 provided detailed feature information, and the “PyRadiomics instruction” included mathematical definitions for each feature [40].

Dosimics data acquisition, preprocessing, and feature extraction were similar to radiomics, but each voxel in the dosimics used dose values instead of grey values (Figs. 1G–I). Before feature extraction, the spatial dose was discretized into a fixed bin width of 1 Gy[RBE]. A total of 93 dosimics features were extracted, as the shape features were identical for both dosimics and radiomics.

Step 2. Feature selection

A bootstrap method was used to draw samples with replacement from the dataset to generate a resampling dataset with the same sample size. We randomly generated 1000 resampling datasets each for the training and test sets. Radiomics and dosimics data were standardized using Z-scores, enhancing comparability across different units and scales.

1) Univariable logistic regression was performed on each feature using 1000 bootstrap datasets generated from the training set. The Wald test p-value was calculated to identify features strongly associated with AOM (median p-value < 0.05). The area under the curve (AUC) represents the model discrimination, with values closer to 1 indicating better performance, while an AUC of 0.5 signifies random prediction. During the initial rough screening, we used a threshold (median AUC > 0.6) slightly above random prediction to exclude weak predictors while retaining as many potentially important features as possible for further analysis [17].

2) The selected features were further refined using least absolute shrinkage and selection operator (LASSO) regression with 5-fold cross-validation, excluding variables with minimal contribution to model prediction. Variables frequently selected (> 50 %) across the 1000 bootstrap datasets from the training set were prioritized for inclusion for further analysis.

3) To prevent overfitting and multicollinearity, bootstrapped stepwise backward variable selection (frequency > 50 %) and Spearman’s correlation test ($|\rho| < 0.7$) were used to optimize the model. During each stepwise backward regression, features were iteratively removed to identify the feature subset that minimized the Akaike Information Criterion (AIC) value, which balances model fit and

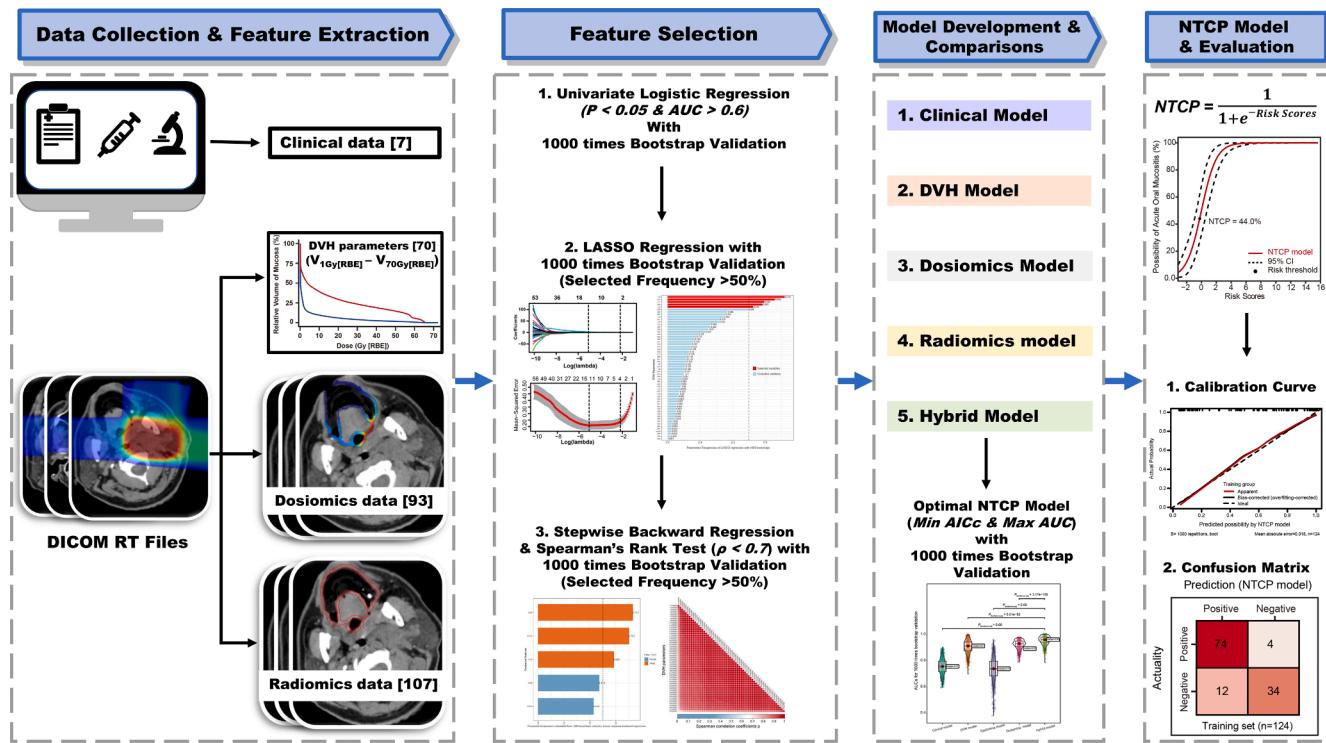


Fig. 2. Overall workflow for developing the normal tissue complication probability (NTCP) model. The numbers in square brackets represent the counts of the variables. Abbreviations: AICc, corrected Akaike information criterion; AUC, Area under the curve; DVH, Dose-volume histogram; LASSO, least absolute shrinkage and selection operator.

complexity. Subsequently, the correlation test was applied to the remaining features, retaining only the variable with the larger AUC for AOM among highly correlated ones.

Step 3. Model development

Logistic regression-based models—clinical, DVH, radiomics, dosimetrics, and hybrid—were developed on the training set using the selected variables. The hybrid model combined features from the single models, refined through stepwise backward regression and Spearman's correlation tests. The model formula is as follows:

$$NTCP = \frac{1}{1 + e^{-(\text{Risk Scores})}}$$

$$\text{Risk Scores} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \dots$$

where β_0 is the intercept, $\beta_1, \beta_2 \dots$ are the coefficients of the features, and $x_1, x_2 \dots$ are the numeric value of the features. The β -values were estimated using maximum likelihood estimation, with 95 % confidence interval (CI) calculated from the bootstrap datasets of the training set. This approach can avoid reliance on the normality assumption required for standard error-based methods and provides relatively stable estimates of variability [17,18,22,41].

Step 4. NTCP model comparison and evaluation

Model performance was assessed using AUC and the corrected AIC (AICc). The AICc is a useful metric for assessing model goodness-of-fit and making model selection, particularly for small sample sizes [16,42]. When comparing candidate models, the smaller the AICc the better the model. Finally, the models with the highest AUC and lowest AICc was deemed the best NTCP model. Generalization performance was evaluated with 1000 bootstrap iterations on the test set, using Mann–Whitney U tests with Bonferroni-adjusted p-values for comparison. Additionally, the AUC difference (ΔAUC) between the hybrid and single models was calculated across 1000 bootstrap validations. A consistently positive ΔAUC indicated the hybrid model's superior

performance.

The accuracy of the NTCP model was assessed using a calibration plot with bootstrap validation on both training and test sets. A calibration curve closer to the identity line indicated a good match between predicted and observed probabilities. A risk threshold for NTCP model was determined to stratify the population using accuracy-based optimization, with confusion matrices used to evaluate model classification performance.

All analyses were performed using R software (version 4.2.2). Statistical significance was set at $p < 0.05$.

Results

This study included 178 patients. The incidence of grade 2–3 AOM during CIRT, based on CTCAE criteria, was 65.7 %, with no occurrences of grade 4 AOM. No significant differences were observed between the features of the training and test sets (all p -values > 0.05 , Table 1).

According to univariable selection criteria ($p < 0.05$, $AUC > 0.6$), we identified 58 DVH, 43 radiomics, 75 dosimetrics, and 1 clinical (tumor location) feature (Table S2–S5). No additional screening for clinical features was required. LASSO regression with 1000 bootstrap iterations selected 5 DVH parameters, 2 radiomics features, and 3 dosimetrics features (Fig. S1–S3). Stepwise backward multivariable regression and Spearman's correlation tests assessed model fit and feature correlations (Table S6 and S7). Due to high inter-correlation among DVH parameters (Fig. S4), only $V_{37 \text{ Gy[RBE]}}$, with the highest AUC (0.927), was included in the DVH model. The radiomics model retained Radiomics_GLCM_MaximalCorrelationCoefficient [MCC] and Radiomics_GLCM_MaximumProbability [MP] ($p = -0.64$, $p < 0.001$; selection frequencies: 86.4 % and 56.7 %). The dosimetrics model included Dosimetrics_GLCM_ClusterProminence and Dosimetrics_GLSZM_High-GrayLevelZoneEmphasis [HGLZE] ($p = 0.47$, $p < 0.001$; selection frequencies: 87.8 % and 77.1 %). The hybrid model was constructed using $V_{37 \text{ Gy[RBE]}}$, Dosimetrics_GLSZM_HGLZE, and

Radiomics_GLCM_MCC after stepwise regression and correlation analyses (Table S6 and S7). Detailed mathematical definitions, interpretations, and representative images for Radiomics_GLCM_MCC and Dosiomics_GLSZM_HGLZE are included in Supplementary File 1.

Table 2 presents the features, performances, and coefficients of the five models. The hybrid model outperformed the others, with a training set AUC of 0.932, a test set median AUC of 0.959, and the lowest AICc (88.2). In the test set, the ΔAUC between the single models and the hybrid model was calculated in each bootstrap dataset. The ΔAUCs for the DVH, radiomics, dosimics, and clinical models were -0.043 (95 % CI: -0.004 to -0.141), -0.221 (95 % CI: -0.090 to -0.406), -0.030 (95 % CI: -0.002 to -0.051), and -0.199 (95 % CI: -0.102 to -0.306), respectively; all were significantly lower than the hybrid model (Bonferroni-adjusted p-values < 0.0001 for each). In the hybrid model, $V_{37\text{ Gy[RBE]}}$ ($\beta = 0.21$, 95 % CI: 0.14 – 0.43 , $p < 0.001$), Dosiomics_GLSZM_HGLZE ($\beta = 0.50$, 95 % CI: 0.07 – 1.04 , $p = 0.028$) and Radiomics_GLCM_MCC ($\beta = -3.52$, 95 % CI = -5.38 to -1.96 , $p = 0.032$) were independent predictors of AOM. Consequently, the hybrid model was identified as the NTCP model for predicting CIRT-associated AOM (Fig. 3A).

The optimal risk threshold for the NTCP model was determined to be 44.0 % (Fig. 3A) and was used for risk stratification. This stratification demonstrated superior classification performance, with an accuracy of 87.1 %, sensitivity of 94.9 %, and specificity of 73.9 % in the training set, and an accuracy of 90.7 %, sensitivity of 97.4 %, and specificity of 73.3

% in the test set (Figs. 3B–D). The calibration plots showed satisfactory agreement between model predictions and actual observations, with the curve being close to the ideal 45° line (Figs. 3E and 3F). The mean absolute error (MAE) of the NTCP model was 0.018 in the training set and 0.043 in the test set, whereas the mean square error (MSE) was 0.05 % and 0.26 %, respectively.

Discussion

Early identification of patients with a high risk of AOM during CIRT is critical for optimizing treatment outcomes and QOL, but predictive models remain scarce. In this study, we developed the first NTCP model for predicting grade ≥ 2 AOM using $V_{37\text{ Gy[RBE]}}$, Dosiomics_GLSZM_HGLZE, and Radiomics_GLCM_MCC based on the OMSC method. The model demonstrated an AUC of 0.932 in the training set and 0.959 in the test set, with risk stratification accuracies of 87.1 % and 90.7 %, respectively.

Oral contouring methods that encompass large volumes of musculature often inaccurately represent the oral mucosa, leading to DVH assessments that primarily reflect muscle tissue rather than mucosa [1,37–39,43], and contributing to the poor performance of many oral mucositis models [1]. Dean et al. suggested that contouring differences can significantly impact dose distribution assessments, especially at moderate doses [38]. In a study of CIRT-induced AOM, although the researchers defined an oral mucosal dose surface model to observe the

Table 2
Parameters and performance of models.

Model	Training set β (95 % CI)	<i>P</i> -value	Test set with 1000 bootstrap validations			<i>P</i> Bonferroni-adj (vs. Hybrid model)	
			AUC	AICc	Median AUC (95 % CI)		
DVH model			0.924	91.9	0.911 (0.796 to 0.983)	-0.043 (-0.004 to -0.141)	< 0.0001
Intercept	-1.68 (-2.66 to -1.13)	< 0.001					
Radiomics model	0.29 (0.20 to 0.51)	< 0.001	0.693	155.3	0.729 (0.540 to 0.880)	-0.221 (-0.090 to -0.406)	< 0.0001
Intercept	5.31 (-0.76 to 12.02)	0.058					
Radiomics_GLCM_MCC	-5.80 (-12.03 to -0.26)	0.026					
Radiomics_GLCM_MP	3.90 (-4.88 to 18.26)	0.461					
Dosiomics model			0.920	90.9	0.924 (0.846 to 0.967)	-0.030 (-0.002 to -0.051)	< 0.0001
Intercept	-3.38 (-6.05 to -2.06)	< 0.001					
Dosiomics_GLSZM_HGLZE	0.55 (0.15 to 1.12)	0.013					
Dosiomics_GLCM_CP	0.30 (0.21 to 0.49)	< 0.001					
Clinical model			0.673	159.6	0.755 (0.645 to 0.857)	-0.199 (-0.102 to -0.306)	< 0.0001
Intercept	0.67 (0.19 to 1.22)	0.009					
Nose cavity vs. Oral cavity	1.41 (-0.10 to 17.83)	0.076					
Nose cavity vs. Pharynx	0.02 (-0.161 to 16.62)	0.977					
Nose cavity vs. Parotid Gland and Ear	-1.36 (-2.65 to -0.30)	0.015					
Nose cavity vs. Other	-0.85 (-2.65 to 0.71)	0.194					
Hybrid model			0.932	88.2	0.959 (0.884 to 1.0)	reference	reference
Intercept	0.07 (-1.23 to 1.37)	0.215					
$V_{37\text{ Gy[RBE]}}$	0.21 (0.14 to 0.43)	< 0.001					
Dosiomics_GLSZM_HGLZE	0.50 (0.07 to 1.04)	0.028					
Radiomics_GLCM_MCC	-3.52 (-5.38 to -1.96)	0.032					

Abbreviations: GLSZM, Gray Level Size Zone Matrix; MP, Maximum Probability; CP, Cluster Prominence; HGLZE, High Gray Level Zone Emphasis; HLCM, Gray Level Co-occurrence Matrix; MCC, Maximal Correlation Coefficient; RBE, relative biological effectiveness, 95 % CI: Confidence interval. AUC, area under the curve; AICc, corrected Akaike Information Criterion

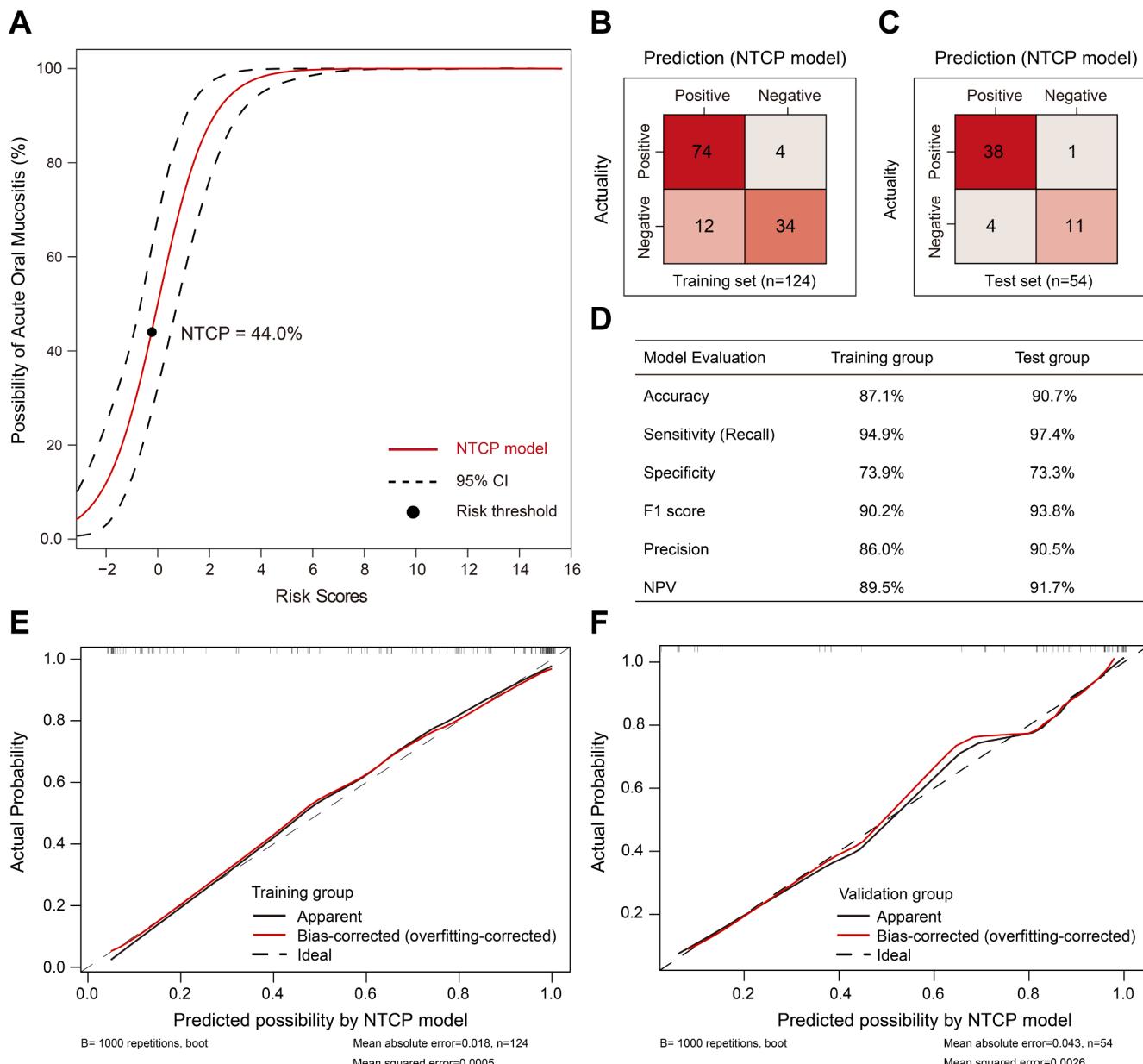


Fig. 3. Multivariable logistic regression normal tissue complication probability (NTCP) model (A). Confusion matrix for risk stratification in the training set (B) and test set (C) and evaluation of its classification performance (D). Calibration curves of the NTCP model with overfitting correction on the training set (E) and test set (F). Abbreviations: CI: Confidence interval; NPV: Negative Predictive Value. (A) Multivariable logistic regression normal tissue complication probability (NTCP) model. (B) Confusion matrix for risk stratification in the training set. (C) Confusion matrix for risk stratification in the test set. (D) Evaluation of classification performance. Abbreviations: NPV: Negative Predictive Value. (E) Calibration curve of the NTCP model with overfitting correction on the training set. (F) Calibration curve of the NTCP model with overfitting correction on the test set.

mucositis area [3], their outlining method was based solely on the volume of the oral cavity, including only the tongue and palate. Conversely, the OMSC method allows for the evaluation of the entire oral mucosa, including typically excluded areas like the buccal mucosa and gingiva. Agheli et al. demonstrated that OMSC-based models could accurately localize relevant parameters, achieving an AUC of 91.7 % and an accuracy of 90.0 % [44]. Therefore, the OMSC may enhance the development of NTCP models by more accurately representing the mucosa's anatomical localization [1,37,38].

The study identified V_{37} Gy[RBE] as the strongest DVH predictor. Given multicollinearity, DVH models are usually univariable [15,45]. In regression models, a regression coefficient represents the average change in the dependent variable for a one-unit increase in the independent variable, assuming all other variables remain constant [45,46].

However, when independent variables are highly correlated, their consistent changes make it challenging for the model to estimate each variable's effect independently, potentially leading to misleading conclusions [46]. DVH parameters, derived from various points of the dose distribution curve, often vary synchronously and convey redundant information [45]. Consequently, adding more DVH parameters to the model does not enhance performance and may instead destabilize the model and undermine its interpretability. In this case, a model with a single DVH parameter is often a more suitable choice [45]. Consistent with previous reports [1,15,16,45], most DVH parameters were significantly associated with AOM in this study, showing similar predictive performance and high inter-correlation (Table S2 and Fig. S4). The middle-to-high dose range is generally recommended for AOM prediction [1,3,9,15,37]. Narayan et al. found that cumulative doses less than

32 Gy were associated with mild acute mucositis, while doses > 39 Gy were associated with a longer duration of mucositis [25]. Other studies highlighted predictors such as $V_{30\text{Gy}}$, $V_{40\text{Gy}}$, and $V_{50\text{Gy}}$ [39,47]. In a CIRT study, $V_{5\text{Gy}} - V_{40\text{Gy}}$ was more influential in severe AOM [3]. In this study, univariable DVH model performance progressively decreased, centering on $V_{37 \text{ Gy[RBE]}}$ (Table S2). Despite minimal differences between medium- and high-dose ranges, the $V_{37 \text{ Gy[RBE]}}$ -DVH model emerged as the best based on clinical relevance and fitting performance. Variations in sample size, variable selection, and contouring methods may account for discrepancies across studies. For instance, while many studies use 5 Gy[RBE] increments [3,15,44], we used 1 Gy[RBE] increments; with coarser increments, $V_{35 \text{ Gy[RBE]}}$ might have been optimal.

Radiomics has been used in NTCP models to capture the heterogeneity of oral mucosal structures [48]. Subtle structural differences in the mucosa across subregions or individuals may lead to varying irradiation-induced injuries [26]. While DVH parameters cannot detect these microstructural differences, radiomics can easily quantify them [27]. In our NTCP model, Radiomics_GLCM_MCC, a measure of texture complexity, was an independent predictor [48]. Lower MCC values, indicating more complex texture and higher gray-level nonuniformity, were associated with an increased risk of AOM. Although the exact mechanisms behind this relationship are unclear, it may involve mucosal nonuniformity due to injury, repair, or patient-specific factors [26]. By quantifying these differences, Radiomics_GLCM_MCC enhances the model's ability to make individualized predictions.

Spatial dosimetrics, an emerging feature in radiotherapy, captures information on spatial dose distribution that conventional DVH parameters miss [17]. Due to variations in spatial dose distribution, individual differences in AOM incidence may arise even among patients with similar DVH parameters [21]. Murakami et al. recommended dosimetric features as a new metric or alternative for evaluating radiotherapy plans [17], and Placidi et al. confirmed their reproducibility and predictive potential, particularly in predictive models [19]. In our NTCP model, Dosiomics_GLSZM_HGLZE emphasized the middle- and high-dose regions, with its weight increasing sharply as the dose rose, underscoring the importance of these dose levels. Importantly, this metric, weakly correlated with $V_{37 \text{ Gy[RBE]}}$ ($\rho = 0.47$, $p < 0.001$), enhanced model performance.

Although tumor location can influence AOM occurrence, it was not included in the final NTCP model. Tumors in the nasal and oral cavities were associated with grade 2–3 AOM, consistent with previous studies [5], but performed poorly in the univariable model compared to DVH, radiomics, and dosimetrics. The proximity of the tumor to the mucosa likely increases irradiation exposure, a factor better captured by dosimetrics or DVH parameters. Additionally, factors such as age, sex, and tumor stage had no significant impact on mucositis in this study, aligning with the controversial findings in the literature [4,5,44]. As a result, these clinical features were excluded from the final model.

Ensuring high model performance while avoiding overfitting was crucial. Due to the advanced nature of CIRT, obtaining suitable external data was challenging [16,49]. To address this, we randomly divided patients into a training set and an independent test set (7:3) and used bootstrap resampling at each modeling step to reduce randomness of the results. We focused on feature relevance, discriminability, and contribution to AOM, removing redundant features to prevent overfitting. In the test set, we conducted 1,000 bootstrap iterations to evaluate and compare the model, demonstrating that the hybrid model consistently outperformed others, with ΔAUC always > 0 ($p < 0.0001$). Although the performance improvement over single DVH and dosimetrics models was modest (median ΔAUC of 0.043 and 0.030, respectively), the hybrid model's ability to capture subtle variations in mucosal structure and spatial dose distribution enhances its individualization. To facilitate clinical use, we classified patients with model-predicted probabilities above 44.0 % as high-risk and those below as low-risk. In clinical practice, when the model identifies high-risk patients, the treatment team can adjust the treatment plan to reduce mucosal damage.

However, for tumors located very close to the mucosa, such adjustments are often difficult to implement. In such cases, early prediction enables timely patient communication and appropriate interventions, helping to ensure the smooth completion of treatment with maintained quality.

This study has some limitations. First, as a retrospective study, selection bias was unavoidable. While fractionated doses showed no significant impact on AOM development (Table S5), DVH parameters above $V_{58 \text{ Gy[RBE]}}$ were not further considered due to their suboptimal statistical performance (median $p > 0.05$ or median AUC < 0.6, Table S2). This result was influenced by data distribution and the stringent statistical criteria applied. Specifically, some patients with grade 2 mucositis received a total dose of only 57.6 Gy[RBE], leaving parameters above $V_{58 \text{ Gy[RBE]}}$ undefined for these patients. When $V_{59 \text{ Gy[RBE]}}$ was used a predictor, its value was 0 %, failing to capture the dose-response relationship for these patients and introducing prediction bias that reduced its statistical relevance. Although this limitation did not affect the overall trend in DVH parameter performance (gradually decreasing from $V_{37 \text{ Gy[RBE]}}$), further research is needed to clarify the predictive value of high-dose parameters in AOM development. Second, although we used an independent test set and bootstrap validation, further validation with prospective data is needed, especially with advances in radiotherapy techniques and AOM management. Third, as noted in previous studies [17,18], planning CT does not account for anatomical changes during radiotherapy, potentially introducing bias. Fourth, while MRI offers radiomic advantages, this study relied on CT due to its convenience and the limitations of MRI in patients with metal implants. Fifth, this study did not analyze Grade 3 AOM separately. On the one hand, the limited number of Grade 3 cases may not ensure reliable model performance when analyzed in isolation. On the other hand, Grade 2 AOM represents an earlier stage where timely prediction and intervention can prevent progression. Therefore, we treated Grade 2 and Grade 3 AOM as equally important endpoints. Future research is expected to conduct a dedicated analysis of Grade 3 AOM. Finally, the NTCP model was developed for CIRT, and its applicability to photon and proton therapies remains unknown.

Conclusions

This study used a multiomics approach to develop and validate the first high-performance NTCP model for predicting grade 2–3 AOM in patients with HN cancer undergoing CIRT. These predictive results may support clinical decision making for individualized oral mucosal management and optimization of treatment plans. As CIRT becomes more widely adopted, further research is needed to facilitate its integration into routine clinical practice.

CRediT authorship contribution statement

Xiangdi Meng: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Zhuojun Ju:** Writing – review & editing, Visualization, Validation, Supervision, Software, Investigation. **Makoto Sakai:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Yang Li:** Writing – review & editing, Formal analysis, Data curation. **Atsushi Masha:** Writing – review & editing, Validation, Methodology. **Nobuteru Kubo:** Writing – review & editing, Supervision. **Hidemasa Kawamura:** Writing – review & editing, Supervision. **Tatsuya Ohno:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.radonc.2025.110709>.

References

- [1] Dean JA, Wong KH, Welsh LC, Jones AB, Schick U, Newbold KL, et al. Normal tissue complication probability (NTCP) modelling using spatial dose metrics and machine learning methods for severe acute oral mucositis resulting from head and neck radiotherapy. *Radiother Oncol* 2016;120:21–7. <https://doi.org/10.1016/j.radonc.2016.05.015>.
- [2] Trotti A. Toxicity in head and neck cancer: A review of trends and issues. *Int J Radiat Oncol Biol Phys* 2000;47:1–12. [https://doi.org/10.1016/s0360-3016\(99\)00558-1](https://doi.org/10.1016/s0360-3016(99)00558-1).
- [3] Musha A, Shimada H, Shirai K, Saitoh J, Yokoo S, Chikamatsu K, et al. Prediction of acute radiation mucositis using an oral mucosal dose surface model in carbon ion radiotherapy for head and neck tumors. *PLoS One* 2015;10:e0141734. <https://doi.org/10.1371/journal.pone.0141734>.
- [4] Moslemi D, Nokhandani AM, Otaghssarai MT, Moghadamnia Y, Kazemi S, Moghadamnia AA. Management of chemo/radiation-induced oral mucositis in patients with head and neck cancer: A review of the current literature. *Radiother Oncol* 2016;120:13–20. <https://doi.org/10.1016/j.radonc.2016.04.001>.
- [5] Vera-Llonch M, Oster G, Hagiwara M, Sonis S. Oral mucositis in patients undergoing radiation treatment for head and neck carcinoma. *Cancer* 2006;106:329–36. <https://doi.org/10.1002/cncr.21622>.
- [6] Ray-Chaudhuri A, Shah K, Porter RJ. The oral management of patients who have received radiotherapy to the head and neck region. *Br Dent J* 2013;214:387–93. <https://doi.org/10.1038/sj.bdj.2013.380>.
- [7] Elting LS, Cooksley CD, Chambers MS, Garden AS. Risk, outcomes, and costs of radiation-induced oral mucositis among patients with head-and-neck malignancies. *Int J Radiat Oncol Biol Phys* 2007;68:1110–20. <https://doi.org/10.1016/j.ijrobp.2007.01.053>.
- [8] Moroney LB, Helios J, Ward EC, Crombie J, Wockner LF, Burns CL, et al. Patterns of dysphagia and acute toxicities in patients with head and neck cancer undergoing helical imrt±concurrent chemotherapy. *Oral Oncol* 2017;64:1–8. <https://doi.org/10.1016/j.oraloncology.2016.11.009>.
- [9] Moshia A, Fukata K, Saitoh JI, Shirai K, Abe T, Mizukami T, et al. Tongue surface model can predict radiation tongue mucositis due to intensity-modulated radiation therapy for head and neck cancer. *Int J Oral Maxillofac Surg* 2020;49:44–50. <https://doi.org/10.1016/j.ijom.2019.06.012>.
- [10] Saitoh JI, Koto M, Demizu Y, Suefuji H, Ohno T, Tsuji H, et al. A multicenter study of carbon-ion radiation therapy for head and neck adenocarcinoma. *Int J Radiat Oncol Biol Phys* 2017;99:442–9. <https://doi.org/10.1016/j.ijrobp.2017.04.032>.
- [11] Mohamad O, Sisir BJ, Saha J, Pompos A, Rahimi A, Story MD, et al. Carbon ion radiotherapy: A review of clinical experiences and preclinical research, with an emphasis on DNA damage/repair. *Cancers (Basel)* 2017;9. <https://doi.org/10.3390/cancers9060066>.
- [12] Mizor JE, Hasegawa A, Jingu K, Takagi R, Bessyo H, Morikawa T, et al. Results of carbon ion radiotherapy for head and neck cancer. *Radiother Oncol* 2012;103:32–7. <https://doi.org/10.1016/j.radonc.2011.12.013>.
- [13] Moshia A, Hirai C, Kitada Y, Tsunoda A, Shimada H, Kubo N, et al. Relationship between oral mucositis and the oral bacterial count in patients with head and neck cancer undergoing carbon ion radiotherapy: A prospective study. *Radiother Oncol* 2022;167:65–71. <https://doi.org/10.1016/j.radonc.2021.12.010>.
- [14] Varnava M, Moshia A, Tashiro M, Kubo N, Okano N, Kawamura H, et al. Dose-volume constraints for head-and-neck cancer in carbon ion radiotherapy: A literature review. *Cancer Med* 2023;12:8267–77. <https://doi.org/10.1002/cam4.5641>.
- [15] Hansen CR, Bertelsen A, Zukauskaite R, Johnsen L, Bernchou U, Thwaites DI, et al. Prediction of radiation-induced mucositis of h&n cancer patients based on a large patient cohort. *Radiother Oncol* 2020;147:15–21. <https://doi.org/10.1016/j.radonc.2020.03.013>.
- [16] Li Y, Sakai M, Tsunoda A, Kubo N, Kitada Y, Kubota Y, et al. Normal tissue complication probability model for acute radiation dermatitis in patients with head and neck cancer treated with carbon ion radiation therapy. *Int J Radiat Oncol Biol Phys* 2022;113:675–84. <https://doi.org/10.1016/j.ijrobp.2022.03.002>.
- [17] Murakami Y, Soyano T, Kozuka T, Ushijima M, Koizumi Y, Miyauchi H, et al. Dose-based radiomic analysis (dosimomics) for intensity modulated radiation therapy in patients with prostate cancer: Correlation between planned dose distribution and biochemical failure. *Int J Radiat Oncol Biol Phys* 2022;112:247–59. <https://doi.org/10.1016/j.ijrobp.2021.07.1714>.
- [18] Zhang Z, Wang Z, Yan M, Yu J, Dekker A, Zhao L, et al. Radiomics and dosimetrics signature from whole lung predicts radiation pneumonitis: A model development study with prospective external validation and decision-curve analysis. *Int J Radiat Oncol Biol Phys* 2023;115:746–58. <https://doi.org/10.1016/j.ijrobp.2022.08.047>.
- [19] Placidi L, Gioscio E, Garibaldi C, Rancati T, Fanizzi A, Maestri D, et al. A multicentre evaluation of dosimetrics features reproducibility, stability and sensitivity. *Cancers (Basel)* 2021;13. <https://doi.org/10.3390/cancers13153835>.
- [20] Placidi L, Lenkowicz J, Cusumano D, Boldrini L, Dinapoli N, Valentini V. Stability of dosimetrics features extraction on grid resolution and algorithm for radiotherapy dose calculation. *Phys Med* 2020;77:30–5. <https://doi.org/10.1016/j.ejmp.2020.07.022>.
- [21] Yang SS, OuYang PY, Guo JG, Cai JJ, Zhang J, Peng QH, et al. Dosimomics risk model for predicting radiation induced temporal lobe injury and guiding individual intensity-modulated radiation therapy. *Int J Radiat Oncol Biol Phys* 2023;115:1291–300. <https://doi.org/10.1016/j.ijrobp.2022.11.036>.
- [22] Rossi L, Bijman R, Schilleman W, Aluwini S, Cavedon C, Witte M, et al. Texture analysis of 3d dose distributions for predictive modelling of toxicity rates in radiotherapy. *Radiother Oncol* 2018;129:548–53. <https://doi.org/10.1016/j.radonc.2018.07.027>.
- [23] Gabryś HS, Buettner F, Sterzing F, Hauswald H, Bangert M. Design and selection of machine learning methods using radiomics and dosimetrics for normal tissue complication probability modeling of xerostomia. *Front Oncol* 2018;8:35. <https://doi.org/10.3389/fonc.2018.00035>.
- [24] Monti S, Xu T, Liao Z, Mohan R, Celli L, Palma G. On the interplay between dosimetrics and genomics in radiation-induced lymphopenia of lung cancer patients. *Radiother Oncol* 2022;167:219–25. <https://doi.org/10.1016/j.radonc.2021.12.038>.
- [25] Narayan S, Lehmann J, Coleman MA, Vaughan A, Yang CC, Enepekides D, et al. Prospective evaluation to establish a dose response for clinical oral mucositis in patients undergoing head-and-neck conformal radiotherapy. *Int J Radiat Oncol Biol Phys* 2008;72:756–62. <https://doi.org/10.1016/j.ijrobp.2008.01.060>.
- [26] Sonis ST. The pathobiology of mucositis. *Nat Rev Cancer* 2004;4:277–84. <https://doi.org/10.1038/nrc1318>.
- [27] Mayerhofer ME, Materka A, Langs G, Häggström I, Szczypinski P, Gibbs P, et al. Introduction to radiomics. *J Nucl Med* 2020;61:488–95. <https://doi.org/10.2967/jnumed.118.222893>.
- [28] 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. <https://doi.org/10.3389/fonc.2020.01708>.
- [29] Vugrin G, Wambwersie A, Koto M, Ohno T, Uhl M, Fossati P, et al. A step towards international prospective trials in carbon ion radiotherapy: Investigation of factors influencing dose distribution in the facilities in operation based on a case of skull base chordoma. *Radiat Oncol* 2019;14:24. <https://doi.org/10.1186/s13014-019-1224-1>.
- [30] Shiba S, Okamoto M, Kiyohara H, Ohno T, Kaminuma T, Asao T, et al. Prospective observational study of high-dose carbon-ion radiotherapy for pelvic recurrence of rectal cancer (gunma 0801). *Front Oncol* 2019;9:702. <https://doi.org/10.3389/fonc.2019.00702>.
- [31] Ohno T, Kanai T, Yamada S, Yusa K, Tashiro M, Shimada H, et al. Carbon ion radiotherapy at the gunma university heavy ion medical center: New facility set-up. *Cancers (Basel)* 2011;3:4046–60. <https://doi.org/10.3390/cancers3044046>.
- [32] Kanai T, Endo M, Minohara S, Miyahara N, Koyama-ito H, Tomura H, et al. Biophysical characteristics of himac clinical irradiation system for heavy-ion radiation therapy. *Int J Radiat Oncol Biol Phys* 1999;44:201–10. [https://doi.org/10.1016/s0360-3016\(98\)00544-6](https://doi.org/10.1016/s0360-3016(98)00544-6).
- [33] Kase Y, Kanai T, Sakama M, Tameshige Y, Himukai T, Nose H, et al. Microdosimetric approach to nirs-defined biological dose measurement for carbon-ion treatment beam. *J Radiat Res* 2011;52:59–68. <https://doi.org/10.1269/jrr.10062>.
- [34] Kubo N, Kubota Y, Kawamura H, Oike T, Sakai M, Kumazawa T, et al. Dosimetric parameters predictive of nasolacrimal duct obstruction after carbon-ion radiotherapy for head and neck carcinoma. *Radiother Oncol* 2019;141:72–7. <https://doi.org/10.1016/j.radonc.2019.07.022>.
- [35] Nachankar A, Moshia A, Kubo N, Kawamura H, Okano N, Sato H, et al. Dosimetric analysis of intraocular hemorrhage in nonsquamous head and neck cancers treated with carbon-ion radiotherapy. *Radiother Oncol* 2022;170:143–50. <https://doi.org/10.1016/j.radonc.2022.02.032>.
- [36] Ueno D, Sato J, Igarashi C, Ikeda S, Morita M, Shimoda S, et al. Accuracy of oral mucosal thickness measurements using spiral computed tomography. *J Periodontol* 2011;82:829–36. <https://doi.org/10.1902/jop.2010.100160>.
- [37] Dean JA, Welsh LC, Wong KH, Aleksić A, Dunne E, Islam MR, et al. Normal tissue complication probability (NTCP) modelling of severe acute mucositis using a novel oral mucosal surface organ at risk. *Clin Oncol (R Coll Radiol)* 2017;29:263–73. <https://doi.org/10.1016/j.clon.2016.12.001>.
- [38] Dean JA, Welsh LC, Gulliford SL, Harrington KJ, Nutting CM. A novel method for delineation of oral mucosa for radiotherapy dose-response studies. *Radiother Oncol* 2015;115:63–6. <https://doi.org/10.1016/j.radonc.2015.02.020>.
- [39] Li K, Yang L, Hu QY, Chen XZ, Chen M, Chen Y. Oral mucosa dose parameters predicting grade ≥3 acute toxicity in locally advanced nasopharyngeal carcinoma patients treated with concurrent intensity-modulated radiation therapy and chemotherapy: An independent validation study comparing oral cavity versus mucosal surface contouring techniques. *Transl Oncol* 2017;10:752–9. <https://doi.org/10.1016/j.tranon.2017.06.011>.
- [40] van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017;77:e104–7. <https://doi.org/10.1158/0008-5472.CAN-17-0339>.
- [41] Kim S, Byun HK, Shin J, Lee IJ, Sung W. Normal tissue complication probability modeling of severe radiation-induced lymphopenia using blood dose for patients with hepatocellular carcinoma. *Int J Radiat Oncol Biol Phys* 2024;119:1011–20. <https://doi.org/10.1016/j.ijrobp.2023.11.060>.
- [42] Portet S. A primer on model selection using the akaike information criterion. *Infect Dis Model* 2020;5:111–28. <https://doi.org/10.1016/j.idm.2019.12.010>.

- [43] 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:S10–9. <https://doi.org/10.1016/j.ijrobp.2009.07.1754>.
- [44] Agheli R, Siavashpour Z, Reiazi R, Azghandi S, Cheraghi S, Paydar R. Predicting severe radiation-induced oral mucositis in head and neck cancer patients using integrated baseline ct radiomic, dosimetry, and clinical features: A machine learning approach. *Heliyon* 2024;10:e24866. <https://doi.org/10.1016/j.heliyon.2024.e24866>.
- [45] Ellsworth SG, van Rossum PSN, Mohan R, Lin SH, Grassberger C, Hobbs B. Declarations of independence: How embedded multicollinearity errors affect dosimetric and other complex analyses in radiation oncology. *Int J Radiat Oncol Biol Phys* 2023;117:1054–62. <https://doi.org/10.1016/j.ijrobp.2023.06.015>.
- [46] Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in regression analyses conducted in epidemiologic studies. *Epidemiology (Sunnyvale)* 2016;6. <https://doi.org/10.4172/2161-1165.1000227>.
- [47] Sanguineti G, Sormani MP, Marur S, Gunn GB, Rao N, Cianchetti M, et al. Effect of radiotherapy and chemotherapy on the risk of mucositis during intensity-modulated radiation therapy for oropharyngeal cancer. *Int J Radiat Oncol Biol Phys* 2012;83:235–42. <https://doi.org/10.1016/j.ijrobp.2011.06.2000>.
- [48] Chiesa-Estomba CM, Mayo-Yanez M, Guntinas-Lichius O, Vander-Poorten V, Takes RP, de Bree R, et al. Radiomics in hypopharyngeal cancer management: A state-of-the-art review. *Biomedicines* 2023;11. <https://doi.org/10.3390/biomedicines11030805>.
- [49] Buizza G, Paganelli C, D'Ippolito E, Fontana G, Molinelli S, Preda L, et al. Radiomics and dosimetrics for predicting local control after carbon-ion radiotherapy in skull-base chordoma. *Cancers (Basel)* 2021;13. <https://doi.org/10.3390/cancers13020339>.