Predicting patient body weight and volumes changes using machine learning for head and neck adaptive radiotherapy

การทำนายการเปลี่ยนแปลงน้ำหนักตัวและปริมาตรของผู้ป่วยมะเร็งศีรษะและลำคอระหว่างการรักษาด้วยการฉายรังสีชนิดปรับแต่งได้

Phawadee Supreeyaporn[1], Chirasak Khamfongkhruea[1], Sasikarn Chamchod[1], Thitiwan Prachanukul[2], Sawanee Suntiwong[2], Todsaporn Fuangrod[1]

[1]Princess Srisavangavadhana College of Medicine, Chulabhorn Royal Academy, Bangkok, Thailand

[2]Radiation Oncology Department, Chulabhorn hospital, Chulabhorn Royal Academy, Bangkok, Thailand

Corresponding author:

Todsaporn Fuangrod

Princess Srisavangavadhana College of Medicine, Chulabhorn Royal Academy, Bangkok, Thailand

906 Khamphaengpet 6 Rd., Lak Si, Bangkok Thailand, 10210

E-mail: todsaporn.fua@cra.ac.th

ภาวดี สุปรียาพร[1], จีรศักดิ์ คำฟองเครือ[1], ศศิกาญจน์ จำจด[1], ฑิติวรรณ ประชานุกูล[2], สวนีย์ สันติวงศ์[2], ทศพร เฟื่องรอด[1]

[1]วิทยาลัยแพทยศาสตร์ศรีสวางควัฒน ราชวิทยาลัยจุฬาภรณ์ กรุงเทพ

[2]งานรังสีรักษามะเร็งวิทยา โรงพยาบาลจุฬาภรณ์ ราชวิทยาลัยจุฬาภรณ์ กรุงเทพ

ผู้นิพนธ์ประสานงาน

ทศพร เฟื่องรอด

วิทยาลัยแพทยศาสตร์ศรีสวางควัฒน ราชวิทยาลัยจุฬาภรณ์

906 ถ.กำแพงเพชร 6 แขวงตลาดบางเขน เขตหลักสี่ กรุงเทพ 10210

อีเมล: todsaporn.fua@cra.ac.th

Submitted: Nov 6, 2022

Revised: Feb 27, 2023

Accepted: Mar 7, 2023

Abstract

Background: Patient anatomical change is the most challenging aspect of head and neck cancer (HNC) patients treated with adaptive radiotherapy (ART) techniques. These can lead to dosimetric deviations and result in an increase in severe radiation toxicity.

Objectives: To use machine learning (ML) in predicting anatomical changes and indicating the optimal time to modify plans in HNC patients for decision-making.

Materials and methods: Volumes of interest (VOI) of 183 Cone-beam computed tomography (CBCT) image datasets were defined based on planning Computed tomography (CT) images. The percentage of body weight (BW), Gross target volume (GTV), Clinical target volume (CTV)70, CTV59.4, Planning target volume (PTV)70, PTV59.4, left parotid gland (PG), and right PG volume changes were retrieved. 143 datasets (78.14%) were applied for training with various ML algorithms; support vector machines (SVMs), kernel approximation regression (KAR), gaussian process regression, Ensembles of trees (ETs), linear regression (LR), regression trees (RTs), and neural networks (NNs). Five and ten-fold cross-validation techniques were applied to select the best prediction model for each specific target. The selected model accuracy for a specific target was tested using the blind testing data set (40 plans or 21.86%) using root mean square error (RMSE) and R-square value (R2).

Results: The selected prediction model (k-fold cross validation, RMSE and R2) were RTs (5-fold, 4.10, 0.49), SVMs (5-fold, 7.40, 0.26), SVMs (5-fold, 5.74, 0.27), GPR (5-fold, 4.88, 0.13), SVMs (5-fold, 5.30, -0.04), ETs (5-fold, 3.19, 0.37), ETs (5-fold, 11.56, 0.57), and RTs (5-fold, 8.31, 0.7), respectively. The system generated the predicted data after the first 11th fractions.

Conclusion: This ML-based patient anatomical change model can provide useful information, which could benefit decision-making in treatment plan modification for HN-ART.

Keywords: Adaptive radiation therapy (ART), Head and neck cancer (HNC), Machine learning (ML), Patient anatomical change

บทคัดย่อ

หลักการและเหตุผล: การเปลี่ยนแปลงทางร่างกายของผู้ป่วยระหว่างการรักษาเป็นความท้าทายในการรักษาผู้ป่วยมะเร็งศีรษะและลำคอด้วยวิธีการฉายรังสีชนิดปรับแต่งได้ การเปลี่ยนแปลงดังกล่าวสามารถนำไปสู่การเปลี่ยนแปลงของปริมาณรังสีในแต่ละอวัยวะ ซึ่งอาจนำไปสู่ผลข้างเคียงรุนแรงจากการฉายรังสีที่เพิ่มขึ้นได้

วัตถุประสงค์: ใช้การเรียนรู้ของเครื่อง (Machine learning) ในการทำนายการเปลี่ยนแปลงทางร่างกายเพื่อการเลือกช่วงเวลาที่เหมาะสมในการปรับแผนการรักษาสำหรับผู้ป่วยมะเร็งศีรษะและลำคอ

วัสดุและวิธีการ: ใช้บริเวณที่สนใจจำนวณ 183 ชุดข้อมูลจากภาพ Cone-beam computed tomography (CBCT) โดยแต่ละอวัยวะถูกกำหนดตามภาพเอกซเรย์คอมพิวเตอร์ (Computed tomography, CT) ของการวางแผนการรักษา แล้วทำการการคำนวณร้อยละความแตกต่างของน้ำหนักตัว (BW), Gross target volume (GTV), Clinical target volume (CTV)70, CTV59.4, Planning target volume (PTV)70, PTV59.4, ต่อมน้ำลายหน้ากกหู (Parotid gland, PG) ข้างซ้าย และ PG ข้างขวา จากนั้นใช้ 143 ชุดข้อมูล (78.14%) เพื่อนำมาให้โปรแกรมเรียนรู้ โดยใช้อัลกอริทึม ดังนี้ support vector machines (SVMs), kernel approximation regression (KAR), gaussian process regression, Ensembles of trees (ETs), linear regression (LR), regression trees (RTs) และ neural networks (NNs) ซึ่งใช้ร่วมกับเทคนิค Five และ ten-fold cross-validation ในการเลือกแบบจำลอง (Model) ที่ดีที่สุดของแต่ละตัวแปร จากนั้นนำแบบจำลองมาทดสอบด้วยชุดข้อมูลทดสอบจำนวน 40 ชุด (21.86%) แล้วทำการประเมินผลด้วยค่า root mean square error (RMSE) และ R-square (R2)

ผลการศึกษา: ผลการทดสอบแบบจำลองที่เหมาะสมของแต่ละตัวแปรมีดังนี้ RTs (5-fold, 4.10, 0.49), SVMs (5-fold, 7.40, 0.26), SVMs (5-fold, 5.74, 0.27), GPR (5-fold, 4.88, 0.13), SVMs (5-fold, 5.30, -0.04), ETs (5-fold, 3.19, 0.37), ETs (5-fold, 11.56, 0.57) และ RTs (5-fold, 8.31, 0.7) ตามลำดับ (ลำดับการแสดงผลได้แก่ k-fold cross validation, RMSE และ R2) จากนั้นระบบทำการสร้างการทำนายหลังจากการฉายแสงครั้งที่ 11

ข้อสรุป: แบบจำลองการเปลี่ยนแปลงทางร่างกายโดยใช้การเรียนรู้ด้วยเครื่องสามารถให้ข้อมูลการเปลี่ยนแปลงทางร่างกายที่เป็นประโยชน์ต่อการช่วยตัดสินใจในการปรับแผนการรักษาของผู้ป่วยมะเร็งศีรษะและลำคอระหว่างการรักษาด้วยการฉายรังสีชนิดปรับแต่งได้

คำสำคัญ: การฉายรังสีชนิดปรับแต่งได้, การเปลี่ยนแปลงทางร่างกายของผู้ป่วย, การเรียนรู้ด้วยเครื่อง, มะเร็งศีรษะและลำคอ

J Thai Assoc Radiat Oncol 2023; 29(1): R67-R88

Introduction

Due to the target volume surrounded by several critical organs at risk (OARs [1], advanced radiotherapy (RT) techniques, such as volumetric modulated arc therapy (VMAT), are widely used for treating head and neck cancer (HNC). This technique allows highly conformal doses to target volumes while reducing toxicity to OARs[2,3]. However, treatment deviations obtain significant anatomical changes that can lead to severe toxicity to OARs. These changes resulted from patient weight loss, tumor location shift, or

shrinkage during the treatment course[4–6]. The toxicity investigation after VMAT treatment revealed that the most prevalent acute toxicities were mucositis, dysphagia, and dermatitis. Meanwhile, late toxicities included xerostomia, laryngeal stiffness, and skin fibrosis[7]. Several studies indicated that anatomical changes impact dosimetric changes in target volumes and OARs (e.g., spinal cord, brainstem, and parotid gland)[6,8–11]. These can result in a tumor underdose or OARs overdose [10,12,13].

There are many approaches to deal with dose deviation from anatomical changes. For example, adaptive radiotherapy (ART) is a modification plan based on anatomical changes. This approach is promising due to the reduction of dosimetric changes and toxicities[10,14]. There are three strategies of ART: offline ART processed before the following fraction of treatment; online ART occurring prior to each treatment fraction; and inline (real-time) ART adapted during treatment delivery. The therapeutic aim, accessible technology, and available resources influence the choice of an ART timescale[15]. However, ART implementation is time-consuming and requires resources because there are four main processes in clinical offline ART workflows: image acquisition, evaluation, replanning, and quality assurance (QA)[10,15–19]. Before treatment delivery, image-guided radiotherapy (IGRT), such as cone-beam computed tomography (CBCT), has been essential for position verification. In addition, some studies have indicated that CBCT is used to investigate dosimetric and volumetric changes through registration planning CT (pCT) with CBCT images[8,10]. Additionally, these images include anatomical change data that may be utilized to identify anatomical trends[8].

According to ART, some cases need to be replanned because of significant anatomical changes. In addition, a study using ART for HNC reported dosimetric benefits of the target volume. Moreover, the doses to target volume were more homogeneous than the non-replanned plans[20]. Anatomical and dosimetric changes during treatment courses are influenced by a number of factors. Numerous researchers have attempted to investigate the variables that indicate the requirement of HNC patients for ART. There have been studies comparing patients with and without replans. According to the findings, the replanned group substantially had a larger ipsilateral parotid gland dosage than the re-CT group without replanning. The dosages to the brainstem and spinal cord were also nearly dramatically raised[21,22]. Another study used deformable image registration (DIR) with Dice Similarity Coefficient (DSC) index, a method for assessing the similarity between two groups, to compare pCT and rescan images to forecast the optimal time between rescanning and ART. In the planning target volume (PTV), brain stem, and body outline, the replanning group substantially had DSC values lower than the non-replanning group ($p < 0.05$). The replanning group had fewer overlapping volumes than the non-replanning group because the reduced DSC value corresponds to the smaller overlapping volume[23]. Research on the use of ART for HNC

also discussed the dosimetric advantages of target volume and OARs. Some studies revealed dose differences of target volume below 1% for the D2% and between 5% and 9% for the D95% and 1.9 Gray (Gy) to gross tumor volume (GTV). In addition, ART helped lower OAR dosages, particularly for the parotid gland and spinal cord [8,10,11]. The dosages reaching the target volume were more homogenous than the non-replanned plans[24]. The recommendation for the frequency of replanning was 1 to 3 times over the treatment course, such as replanning at the first, second, fifth week, or after the fourth week[20,25]. However, CBCT imaging requires more acquisition time and provides more accumulated imaging dose than 2-dimensional orthogonal kV imaging[26,27]. Therefore, trends of predicting anatomical change, including weight change and volumetric change, can assist in quantifying the optimized replanning time.

Machine learning is moving forward in the radiation therapy field. Applications include image fusion, auto-segmentation, planning optimization, treatment outcome, toxicity predictions, QA, ART, and clinical decision support systems[28–33]. Models for predicting anatomical change in HNC treatment have focused on weight loss and volumetric tumor change. The methods, which were utilized to create predictive models to analyze anatomical change over treatment courses, included curve fitting tools (exponential and linear models), logistic regression, and classification and regression tree (CART) analysis[34–37]. For weight loss prediction models, logistic regression was applied to determine the major factors of crucial weight loss. The significant factors were concurrent chemoradiotherapy (CCRT) status, treatment site, prescription dosage, RT on the neck, and RT method[35,36].

The pCT and CBCT with OAR segmentations were registered to quantify the difference between tumor and OAR volumes, including GTV, clinical target volume (CTVs), PTVs, and the parotid gland (PG). The absolute and percentage differences of GTV and PG were acquired and evaluated. Curve fitting with exponential and linear models was used to fit the GTV and PGs volume data across fractions (treatment days). According to the results, GTV change derived from the exponential model while PG change fitted to the linear model. The R-squared values of both models were > 0.9 for both GTV and PG changes[34]. Displacement field (DF) between pCT and CBCT using B-spline and Demon DIR algorithms can be applied to evaluate the significant patient anatomical change based on DSC[37]. Instead of tracking the patient anatomical change, DF was applied to estimate the dose distribution change based on pCT and CBCT at specific treatment fractions towards the development of ART trigger[37].

This study aims to develop the prediction model of target volume change using various machine learning algorithms to determine the optimal time for replanning in HN ART after 11 fractions of treatment delivery. Because most cases in our center underwent CBCT image acquisitions at the 1st, 11th, 16th, and 21st fractions, there were only 1 to 2 CBCT image

datasets within the second week of treatment. Therefore, we used the image dataset between the 1st and 11th fractions to represent trend change. The optimal time point can assist staff in replan decision-making and/or the frequency of CBCT image acquisition.

Materials and methods

Patient selection and data acquisition

This study was authorized by the Human Research Ethics Committee of Chulabhorn Research Institute, Thailand (026/2565, Apr 22 2022). This research is a retrospective analysis of anonymized radiological and clinical data in the Radiation Oncology department at Chulabhorn Hospital, Bangkok, Thailand. One hundred eighty-three datasets from 25 diagnosed patients with primary HNC were included. These patients were treated with RT or CCRT for HNC. Planning CT images, organ contours, and daily CBCT images of HNC patients were collected from 2018 to 2021. All cases had curative aims, had CBCT images greater than or equal to 5 series, and were treated with VMAT using a simultaneous integrated boost technique. In addition, the clinical data, including tumor staging, age, prescription dose for radiotherapy, and concurrent chemotherapy status, were recruited for analysis. The summary of patient demographic is summarized in **Table 1**.

Treatment planning system

All treatment planning processes took within a week after the CT-sim scan with a supine position immobilized with a thermoplastic mask. Radiation oncologists delineated the tumors. CTV and PTV were defined as extending 5 mm, and GTV was 10 mm extension. The prescription dose to PTV was 66-70 Gy in 33 fractions, and the dose per fraction was 2 – 2.121 Gy. The radiation dose calculations were computed by the Eclipse treatment planning system version 16.1 (Varian Medical Systems, Palo Alto, CA).

Auto-contouring of CBCT images and their verification

Before each treatment fraction, all patients underwent position verification before treatment delivery using CBCT. In our center, CBCT image acquisitions were usually acquired at the 1st, 11th, 16th, and 21st fractions. In addition, some patients underwent CBCT acquisition more frequently than required by this protocol due to the rapid tumor volume change. With the axial length of full fan mode, the field of view (FOV) of CBCT images obtained TrueBeam (Varian Medical System, Palo Alto, CA) was 28 cm in diameter. The isocenter of the CBCT image was in the same position as the treatment plan. The CBCT dataset of each patient was transferred to MIM (MIM Software, Cleveland, Ohio, USA), a commercial software used for Digital Imaging and Communications in Medicine (DICOM) image registration, manipulation, and reading. MIM was used for auto-segmentation into the GTV, CTV70, CTV59.4, PTV70, PTV59.4, and left and right parotid glands. Then, the contours were verified and amended by two radiation oncologists. After that, the volumes outside the FOV of CBCT images were excluded from this study. This means that only the volumes within the FOV were investigated.

Table 1. Summary of patient demographics and characteristics.

Characteristics	Detail
Gender (# patient)	
Male	17
Female	8
Median age (years)	58 (28 – 86)
Diagnosis (# patient)	
Nasopharynx	10
Oropharynx	9
Oral cavity	6
T classification (# patient)	
1	3
2	7
3	8
4	7
N classification (# patient)	
0	4
1	6
2	12
3	3
M Classification (# patient)	
0	25
Treatment method (# patient)	
Radiotherapy alone	2
Concurrent chemoradiotherapy	23
Median prescription dose (Gy)	70 (66 – 70)

Machine learning model

The model construction workflow is outlined in **Figure 1**. A random sampling technique was applied in this study. All the datasets (183 sets) were divided into 143 datasets (20 patients) for the model training and 40 datasets (5 patients) for the model testing. The five patients from the testing dataset included two cases of nasopharyngeal cancer, two cases of oropharyngeal cancer, and one case of oral cancer. These five cases had at least two CBCT datasets within the first 11th fraction for training and the datasets after the 11th fraction for testing the accuracy of the models. In this study, seven different algorithms were used, including Linear regression models (LR), Regression trees (RTs), Support vector machines (SVM), Gaussian process regression models (GPR), Kernel Approximation Regression Models (KAR), Ensembles of trees (ETs), and Neural Networks (NNs). To train the model, 17 features were extracted consisting of age, T classification, N classification, M classification , gender, treatment site, concurrent chemotherapy status, fraction, day difference from planning date, the initial value of patient weight, GTV volume, CTV70 volume, CTV59.4 volume, PTV70 volume, PTV59.4 volume, left PG, and right PG. A feature ranking for regression using the F-tests function for machine learning calculates the f-test score. This score is defined as -ln (p-value). From this equation, the score of the features must be more than 2.996 to fulfil the significance level of 0.05. The response variable was the percentage difference between the initial value and the value in that fraction of body weight (BW), GTV, CTV70, CTV59.4, PTV70, PTV59.4, left PG, and right PG. These features of 143 training data sets were entered into the Regression Learner of MATLAB R2022a (Mathworks Inc., Natick, MA, USA) with 5- and 10-fold cross-validation to improve the performance of the models. After training, the results were compared to find the best model for each output using RMSE, and R^2. Then, the best-performing model was selected and extracted into functions.

Model performance evaluation using blind tests

All the models were tested with the independent 40 datasets in terms of RMSE, and R^2 using the following equations:

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(X_i - Y_i)^2} \quad (1)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(X_i - Y_i)^2}{\sum_{i=1}^{m}(\overline{Y} - Y_i)^2} \quad (2)$$

Where m is the observation number, Xi is the predicted value, Yi is the actual value, and is the average of the actual values.

After that, the best performance models of body weight and all volumes were extracted and generated to be functions. These functions were trained with the 143 training datasets and the independent datasets within the first 11th fraction. Then, the trends of anatomical changes were illustrated in line graphs

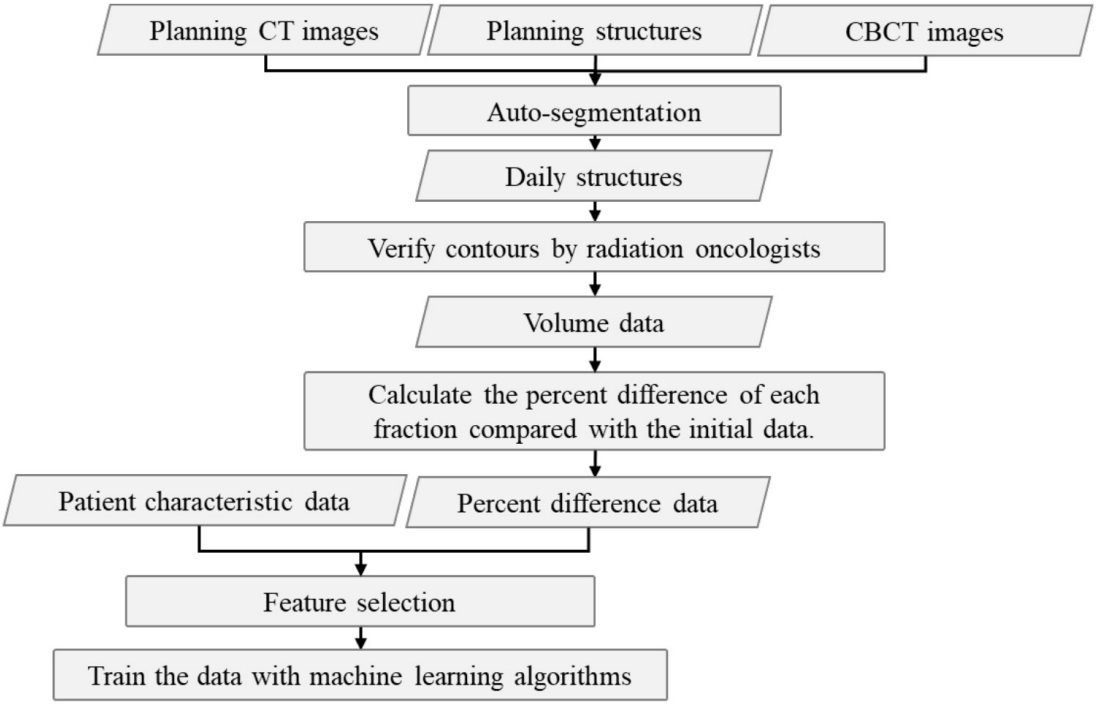

Figure 1. Model construction workflow.

Results

Patient selection

Figure 2 illustrates the percentage differences between the initial and final values of the body weight and volume. The results showed that the average values for body weight, GTV, CTV70, CTV59.4, PTV70, PTV59.4, left and right PG were -6.58 (range: -16.40 to 1.96), -8.16 (range: -42.72 to 26.41), -7.30 (range: -37.14 to 28.92), -4.28 (range: -38.25 to 14.04), -7.09 (range: -31.80 to 20.33), -4.04 (range: -30.99 to 13.13), -21.25 (rang: -57.71 to 9.47), and -15.96 (range: -43.43 to 18.07), respectively. The outliers were found in the GTV, CTV, and PTV in areas below -40 and above 60.

Model construction

Before the training model, the features were ranked by F-test score. Table shows the F-test score ranking of BW, GTV, CTV70, CTV59.4, PTV70, PTV59.4, and left and right PG models. The feature selection eliminated the predictors with an f-test score of less than 2.996. Thus, the features selected for body weight, GTV, CTV70, CTV59.4, PTV70, PTV59.4, left and right PG had 12, 13, 13, 11, 13, 11, 12, and 12 features, respectively. The feature with an F-test score of more than 2.996 was chosen for training. For example, the total of 12 features consists of initial PTV70 volume, initial weight, initial GTV volume, age,

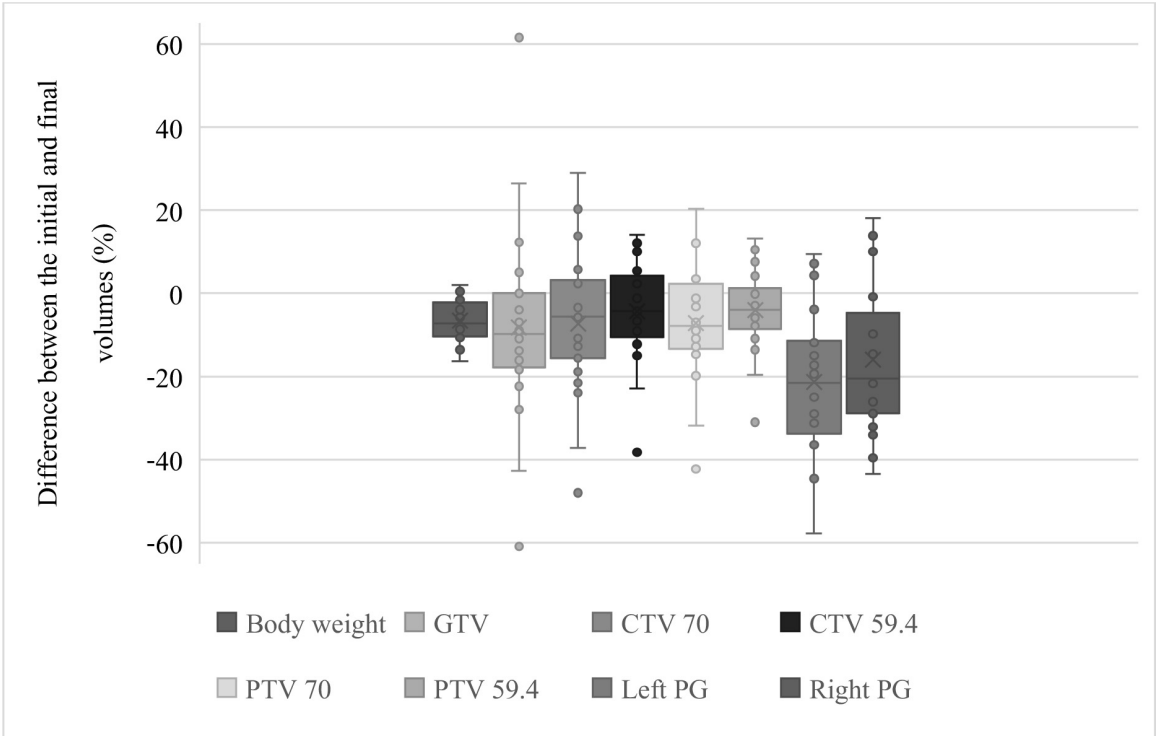

Figure 2. Boxplot of the percentage differences between the initial and final volume of the targets and OAR.

initial PTV59.4 volume, initial CTV70 volume, initial left PG volume, initial CTV59.4 volume, day difference from planning date, fraction, T classification, and initial right PG volume were utilized for BW model training. After that, the selected features were fed into the models for training. The prediction model results of GTV, CTV70, PTV70, BW, left PG, right PG, CTV59.4, and PTV59.4 are shown in **Table 2**, respectively. Moreover, there was no significant difference between the 5-fold and 10-fold cross-validation. Therefore, **Table 2-3** indicates only the performance of the 5-fold cross-validation technique. From these results, the best models of each response variable (k-fold cross validation, RMSE, R2) were RTs (5-fold, 4.10, 0.49) for body weight, SVMs (5-fold, 7.40, 0.26) for GTV, SVMs (5-fold, 5.74, 0.27) for CTV70, GPR (5-fold, 4.88, 0.13) for CTV59.4, SVMs (5-fold, 5.30, -0.04) for PTV70, ETs (5-fold, 3.19, 0.37) for PTV59.4, ETs (5-fold, 11.56, 0.57) for left PG, and RTs (5-fold, 8.31, 0.7) for right PG. These models were generated into functions and evaluated with blind testing.

Table 2. Results of BW, GTV, CTV70, CTV59.4, PTV70, PTV59.4, left and right PG prediction models.

Models	BW		GTV		CTV70		CTV59.4		PTV70		PTV59.4		Left PG		Right PG	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
Linear regression	6.26	-0.24	17.22	-3.01	8.66	-0.67	6.95	-0.75	8.21	-1.49	5.33	-0.95	16.79	0.02	14.78	-0.08
Regression trees	4.10	0.49	13.55	-1.27	11.00	-1.44	8.92	-1.76	11.59	-3.38	8.10	-3.08	12.23	0.52	8.31	0.70
Support vector machines	4.81	0.26	7.40	0.26	5.74	0.27	5.04	0.08	5.30	-0.04	3.70	0.06	12.02	0.50	9.81	0.52
Gaussian process regression	5.05	0.19	11.67	-0.84	6.58	0.04	4.88	0.13	6.04	-0.35	3.65	0.08	15.35	0.18	11.94	0.29
Kernel approximation regression	5.87	-0.10	11.18	-0.69	7.41	-0.22	5.54	-0.11	6.23	-0.44	4.63	-0.47	20.25	-0.43	20.29	-1.04
Ensembles of trees	4.39	0.41	11.09	-0.52	8.53	-0.47	7.64	-1.02	5.94	-0.15	3.19	0.37	11.56	0.57	8.74	0.66
Neural Networks	4.96	0.22	25.89	-8.06	11.24	-1.82	8.30	-1.50	9.92	-2.64	6.68	-2.06	26.97	-1.53	24.15	-1.89

Model evaluation

After training the data, blind independent testing cases were performed for model evaluation. The testing datasets and the first 11th fraction data were trained with the best model of each response variable to predict the anatomical changes of the 12th fraction until the end of treatment. The results of the five testing cases are shown in Table . The average RMSEs of body weight, GTV, CTV70, CTV59.4, PTV70, PTV59.4, and left and right PG were 3.74, 4.23, 3.80, 5.83, 5.00, 3.52, 11.65, 10.29, respectively. While the R^2 were -4.06, 0.53, 0.36, -3.62, -0.76, -0.69, -0.87, 0.19, respectively.

Table 3. Test results of the model prediction with blind independent cases.

Response variables	Patient no.1		Patient no.2		Patient no.3		Patient no.4		Patient no.5	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
Weight	2.55	0.70	2.74	-1.94	3.74	0.67	3.14	-0.59	6.52	-19.15
GTV	2.72	0.69	9.11	0.28	3.55	0.74	3.73	0.42	2.03	0.52
CTV70	2.26	0.63	7.72	0.32	5.19	0.37	1.42	0.90	2.40	-0.45
CTV59.4	0.45	0.97	6.28	-1.52	15.11	-4.68	2.04	0.46	5.26	-13.33
PTV70	4.46	-1.20	6.66	-0.66	9.73	-1.90	2.14	0.59	2.00	-0.62
PTV59.4	3.41	-0.52	1.13	0.66	10.45	-3.57	1.51	0.38	1.12	-0.39
Lt. PG	9.58	-0.53	8.01	0.31	13.82	0.03	6.32	0.77	20.51	-4.93
Rt. PG	9.48	0.10	7.82	0.43	16.08	-0.17	15.01	-0.02	3.06	0.63

abbreviation: GTV = gross tumer volume, CTV = clinical target volume, PTV = planning target volume, PG = paroted gland, LT = left, RT = right

Discussion

For this study, there were some problems with the auto-segmentation of MIM software as follows: There was a significant error in adding a margin to target volumes, particularly near the shoulder region. In addition, the overestimated CTV auto-segmentations were also detected in the other studies[38–40]. Furthermore, CBCT images of each fraction were not scanned at the same position exactly. This situation caused the lower part of the PTV to have many variations in contouring. Moreover, the PTV59.4 volumes were usually more extensive than the FOV of CBCT scanning. As a result, only the parts of the PTV59.4 volumes within the FOV of CBCT scanning could be analyzed.

From the results of this study, the body weight, CTV59.4, and PTV 59.4 prediction

models provided quite desirable performance (RMSE < 5%, and R2 > 0.12). Since almost all the data for training was non-linear and some datasets had high variations, the RTs, SVMs, ETs, and GPR algorithms were appropriate for these response changes[41,42].

However, there were terrible performance models, including left and right PG volumes with RMSE > 8%. Because these volume sizes were too small, minor errors in prediction could result in a large RMSE. In addition, another study used exponential and linear curve fitting to predict GTV and PG shrinkages. The results showed that the exponential curve was proper with GTV and linear regression was optimal with PG with R^2 of more than 0.9[34]. Moreover, as shown in Figure 2, there were significant variations in the final volumes. In addition, some target volumes increased in size at the end of treatment. These occurrences were probably multifactorial and may be related to edema, inflammation, physiology, and dynamic cell processes[43]. The large variations and small sample size caused training machine learning models more difficult and resulted in high errors.

Recently, there were no exact criteria to replan for head and neck cancer. The trends of anatomical change can be a part of the decision support to generate the trigger point for replanning. There is a relative paucity of studies investigating the criteria to make ART. Most of these studies analyzed anatomical changes from weight loss, tumor shrinkage, and parotid gland volume change. The comparison between planning CT and rescanning CT was 9.5%, 5.3%, 8.4%, 20.5%, and 20.9% for GTV, CTV of primary tumors (CTV1), CTV of nodal diseases (CTV2), left PG, and right PG, respectively[44]. In addition, a study indicated that the CTV1 and CTV2 volumes decreased in week two by 3.2% and 10%, respectively. These findings resulted in a 2 Gy and 3.9 Gy decrease in the minimum doses of PTV1 and PTV2, respectively. Furthermore, these volume reductions may result from chemotherapy induction[45]. Therefore, adjusting the model to predict dose change can increase the impact on clinical workflow.

In this study, new datasets from the 1st to the 11th fractions were fed into the generated model function along with the training data set. Then, the model predicted the percentage volume difference of the GTV, CTVs, and PTVs, which starts at the 12th fraction to the end of treatment. **Figure 3** shows an example of the percentage difference of prediction models between the initial and each fraction response variables. The x-axis indicates the fraction of the treatment. While the y-axis indicates the difference in these response variables in percentage. The legend is shown at the bottom left corner, including the actual and predicted values of body weight, GTV, CTVs, PTVs, and PGs. The trained line shows that the actual values are in this fraction, which is the 11th fraction. The blue zone represents the actual values at the 1st to 11th fraction. The yellow zone represents the predicted values at the 12th fraction until the end of the treatment course. In addition, the plot shows the warning point for each variable, consisting of the significant

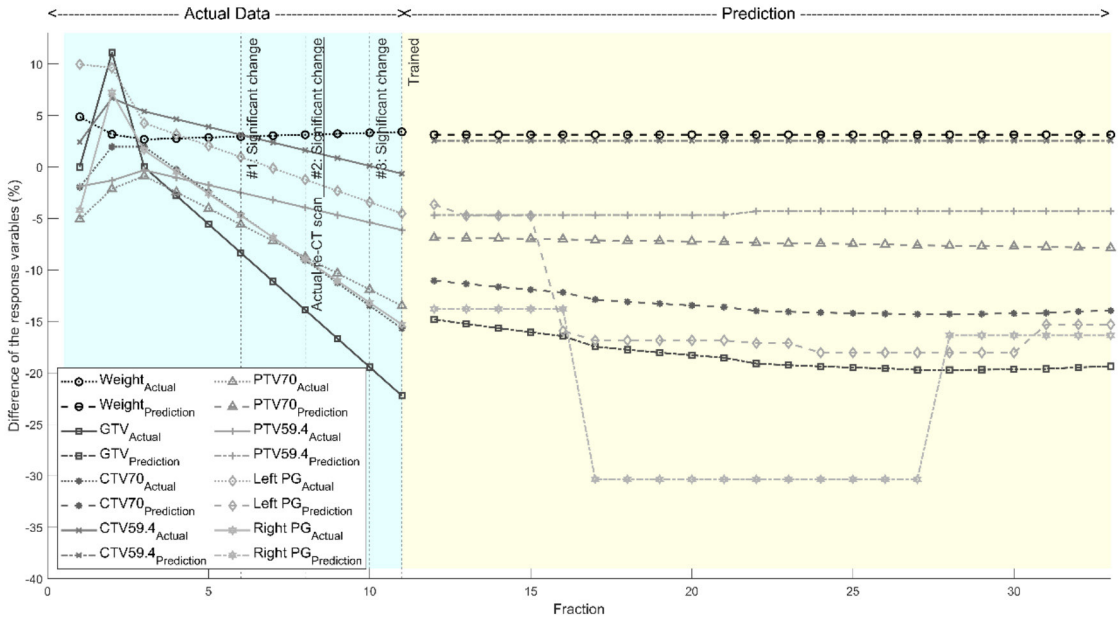

Figure 3. The example of using the percentage difference prediction models.

change point and the CBCT recommendation point. The cut points were -12.24%, -9.5%, -3.2%, -10%, and -26 for BW, GTV, CTV70, CTV59.4, and PGs, respectively[44–46].

In this case, the first CBCT dataset was acquired after the patients underwent CT simulation for 14 days. As shown in the first week of treatment, the GTV volume increased by about 10%. The Schwartz formula can describe this situation: $DT = [\ln 2 \times \Delta T]/[\ln(X_2/X_1)]$. Where DT is the doubling time of the tumor volume (days), ΔT is the time between the two scans (days), and X_1 and X_2 are the initial and final volumes, respectively[47]. Additionally, from a study of tumor progression, they found that the average DT of HNC was 99 days[48]. Thus, the ΔT was 14 days for this case, which resulted in the X_1/X_2 value being 1.103 according to the Schwartz formula. This means the tumor volume may increase the size by about 10% after a waiting time of 14 days.

Moreover, the graph shows that predicted tumor volume change trends were usually higher than the actual values before the 12th fraction. The reason is that the values of the training dataset impacted the predicted values. If most cases had higher values than this example case, the predicted values might be higher than the actual values. It can be seen that the body weight trend after the 11th fraction was constant,

and the algorithm used for this prediction was RTs. Moreover, the ranks of predictors in **Table 1S** show that the day difference from the planning date and fraction was in the 9th and 10th orders, respectively. Therefore, these predictors impact the calculation less than the others. Because the other predictors were the same, the results frequently showed constant values. However, the other graphs of the other cases showed the body weight change more obviously than the graph of this case. In addition, the PG predictions suddenly decreased during the fourth week of treatment. These reductions can result from ongoing reductions between the third and fourth weeks. Furthermore, the right PG prediction indicated a more rapid decline than the left PG because the algorithms used for the right and left PGs were RTs and ETs, respectively. ETs produced smoother results than RTs, which frequently change abruptly after the input is in the conditions of the other branch of the RTs

The limitations of this study are the accuracy and precision of the target volumes, which were affected by the quality of CBCT images and the auto-contouring software. Furthermore, the data set was too small and had high variation, which caused the training results to provide high error. Moreover, after the model can predict dose change and dose-volume histograms (DVHs) and the warning cut-off points of each organ are defined, these can develop a clinical decision support system (DSS). The DSS will provide the data trends and trigger points for staff to make their decision.

Conclusion

This model development from machine learning was a pilot approach for developing anatomical change prediction. The trends of body weight, GTVs, CTVs, PTVs, and PG volume changes for HNC patients can be predicted. According to Table , the model of BW, CTV59.4, and PTV59.4 provided the predicted values that resemble the actual values. If the model accuracy can be improved, the models would be useful for providing decision support before treatment delivery in the following fraction. Further studies on volume changes in other OARs and dosimetric changes in targets and OARs should be investigated. This establishes the tool as a decision support system to trigger the optimal time to replan.

Acknowledgements

The authors would like to thank the lecturers and medical physicists in the Princess Srisavangavadhana College of Medicine, Chulabhorn Royal Academy and Radiation Oncology department, Chulabhorn Hospital for their useful comments and support with this research. 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.

Table 1S F-test results of BW, GTV, CTV70, CTV59.4, PTV70, PTV59.4, left and right PG models.

Rank	BW		GTV		CTV70		CTV59.4		PTV70		PTV59.4		Left PG		Right PG	
	Features	F-test score	Features	F-test score	Features	F-test score	Features	F-test score	Features	F-test score	Features	F-test score	Features	F-test score	Features	F-test score
1	Initial PTV70 volume	26.98	M classification	57.47	M classification	53.62	M classification	59.11	M classification	53.52	Initial classification	56.26	Fraction	30.65	Fraction	20.12
2	Initial weight	14.92	Initial right PG volume	50.30	Initial right PG volume	47.35	Initial PTV70 volume	58.72	Initial right PG volume	47.60	Age	51.33	Day difference from planning date	25.68	Initial right PG volume	20.06
3	Initial GTV volume	14.57	Initial PTV59.4 volume	47.40	Initial left PG volume	45.79	Initial right PG volume	58.62	Age	44.22	Initial weight	50.22	Initial PTV59.4 volume	21.32	Day difference from planning date	19.25
4	Age	12.30	Initial PTV70 volume	46.55	Initial PTV59.4 volume	45.76	Initial weight	58.57	Initial weight	41.38	M classification	47.67	Initial CTV70 volume	13.72	Initial weight	17.08
5	Initial PTV59.4 volume	12.12	Age	44.75	Initial PTV70 volume	43.23	Initial PTV59.4 volume	53.58	Initial PTV59.4 volume	41.06	Initial PTV59.4 volume	46.41	Initial CTV59.4 volume	10.44	Initial left PG volume	16.55
6	Initial CTV70 volume	10.88	Initial left PG volume	43.32	Age	35.42	Initial CTV70 volume	48.42	Initial left PG volume	40.95	Initial PTV70 volume	41.15	Initial left PG volume	10.18	Initial PTV59.4 volume	15.86
7	Initial left PG volume	10.53	Initial GTV volume	41.02	Initial weight	33.26	Age	43.28	Initial PTV70 volume	39.69	Initial CTV70 volume	39.43	Initial PTV70 volume	9.03	Initial PTV70 volume	15.22
8	Initial CTV59.4 volume	9.42	Initial CTV70 volume	39.94	Initial CTV70 volume	32.66	Initial CTV59.4 volume	41.74	Initial CTV70 volume	35.15	Initial CTV59.4 volume	34.23	Initial right PG volume	7.16	Age	11.99

Rank	BW		GTV		CTV70		CTV59.4		PTV70		PTV59.4		Left PG		Right PG	
	Features	F-test score	Features	F-test score	Features	F-test score	Features	F-test score	Features	F-test score	Features	F-test score	Features	F-test score	Features	F-test score
9	Day difference from planning date	9.36	Initial CTV59.4 volume	35.40	Initial GTV volume	29.97	Initial GTV volume	39.03	Initial CTV59.4 volume	33.41	Initial GTV volume	30.93	Initial GTV volume	7.05	Initial CTV59.4 volume	11.49
10	Fraction	9.00	Initial weight	30.36	Initial CTV59.4 volume	29.09	Initial left PG volume	33.29	Initial GTV volume	29.62	Initial left PG volume	30.88	Age	6.39	Initial GTV volume	11.20
11	T classification	6.31	Gender	6.46	Day difference from planning date	5.52	Day difference from planning date	6.98	Day difference from planning date	5.98	Day difference from planning date	7.44	Initial weight	6.26	Initial CTV70 volume	9.65
12	Initial right PG volume	4.97	Treatment sites	5.10	Gender	3.70	Gender	2.77	T classification	3.22	T classification	1.81	Concurrence chemotherapy status	4.16	M classification	5.64
13	Treatment sites	2.44	Day difference from planning date	5.07	Treatment sites	3.61	Concurrence chemotherapy status	2.15	Gender	3.01	Fraction	1.79	Treatment sites	2.94	T classification	1.99
14	Concurrence chemotherapy status	1.38	T classification	2.57	Concurrence chemotherapy status	1.92	T classification	2.06	Treatment sites	2.66	Concurrence chemotherapy status	1.07	M classification	2.30	Concurrence chemotherapy status	1.00
15	M classification	0.11	Fraction	1.50	Fraction	1.56	Treatment sites	2.04	Fraction	1.68	Treatment sites	0.75	T classification	0.68	Treatment sites	0.52
16	Gender	0.04	Concurrence chemotherapy status	1.00	T classification	1.27	Fraction	0.92	Concurrence chemotherapy status	0.05	Gender	0.62	Gender	0.32	Gender	0.39
17	N classification	0.00	N classification	0.00	N classification	0.00	N classification	0.00	N classification	0.00	N classification	0.00	N classification	0.00	N classification	0.00

References

1. Arora A, Purohit R, chigurupalli K, Bhandari M, Gupta A, Peter S. Comparison of Sequential Boost and Simultaneous Integrated Boost Volumetric Modulated Arc Therapy in Treatment of Head and Neck Carcinoma: A Prospective Interventional Study. JCDR 2022; 16: 1-5.
2. Dahele M, Tol JP, Vergeer MR, Jansen F, Lissenberg-Witte BI, Leemans CR, et al. Is the introduction of more advanced radiotherapy techniques for locally-advanced head and neck cancer associated with improved quality of life and reduced symptom burden? Radiother Oncol 2020;151:298–303.
3. Cilla S, Deodato F, Macchia G, Digesù C, Ianiro A, Piermattei A, et al. Volumetric modulated arc therapy (VMAT) and simultaneous integrated boost in head-and-neck cancer: is there a place for critical swallowing structures dose sparing? Br J Radiol 2016;89: 1-10.
4. Gros SAA, Xu W, Roeske JC, Choi M, Emami B, Surucu M. A novel surrogate to identify anatomical changes during radiotherapy of head and neck cancer patients. Med Phys 2017;44:924–34.
5. Noble DJ, Yeap P-L, Seah SYK, Harrison K, Shelley LEA, Romanchikova M, et al. Anatomical change during radiotherapy for head and neck cancer, and its effect on delivered dose to the spinal cord. Radiother Oncol 2019;130:32–8.
6. Rozendaal RA, Mijnheer BJ, Hamming-Vrieze O, Mans A, van Herk M. Impact of daily anatomical changes on EPID-based in vivo dosimetry of VMAT treatments of head-and-neck cancer. Radiother Oncol 2015;116:70–4.
7. Moncharmont C, Vallard A, Mengue Ndong S, Guy J-B, Saget C, Méry B, et al. Real-life assessment of Volumetric Modulated Arc Therapy (VMAT) toxicity in Head and Neck Squamous Cell Carcinoma (HNSCC) treatment. Acta Oto-Laryngologica 2016;136:181–8.
8. Wang H, Xue J, Chen T, Qu T, Barbee D, Tam M, et al. Adaptive radiotherapy based on statistical process control for oropharyngeal cancer. J Appl Clin Med Phys 2020;21:171–7.
9. Sun L, Kirkby C, Smith W. Dosimetric effect of body contour changes for prostate and head and neck volumetric modulated arc therapy plans. J Appl Clin Med Phys 2019;20:115–24.
10. Belshaw L, Agnew CE, Irvine DM, Rooney KP, McGarry CK. Adaptive radiotherapy for head and neck cancer reduces the requirement for rescans during treatment due to spinal cord dose. Radiat Oncol 2019;14:189-195.
11. Chibane BI, Benrachi F, Salah Bali M. Adaptive approach for nasopharyngeal

carcinoma patients during Volumetric Modulated Arc Therapy treatment (VMAT). Int J Radiat Res 2020;18:369–74.
12. van Beek S, Jonker M, Hamming-Vrieze O, Al-Mamgani A, Navran A, Remeijer P, et al. Protocolised way to cope with anatomical changes in head & neck cancer during the course of radiotherapy. Tech Innov Patient Support Radiat Oncol 2019;12:34–40.
13. Alexander KM, Gooding J, Schreiner LJ, Olding T. Clinical management of tumour volume changes in VMAT head & neck radiation treatment. J Phys: Conf Ser 2017;847:1-5.
14. Navran A, Heemsbergen W, Janssen T, Hamming-Vrieze O, Jonker M, Zuur C, et al. The impact of margin reduction on outcome and toxicity in head and neck cancer patients treated with image-guided volumetric modulated arc therapy (VMAT). Radiother Oncol 2019;130:25–31.
15. Green OL, Henke LE, Hugo GD. Practical Clinical Workflows for Online and Offline Adaptive Radiation Therapy. Semin Radiat Oncol 2019;29:219–27.
16. Gros SAA, Santhanam AP, Block AM, Emami B, Lee BH, Joyce C. Retrospective clinical evaluation of a decision-support software for adaptive radiotherapy of Head & Neck cancer patients. Front Oncol 2022;12:1-20.
17. Lim SB, Tsai CJ, Yu Y, Greer P, Fuangrod T, Hwang K, et al. Investigation of a Novel Decision Support Metric for Head and Neck Adaptive Radiation Therapy Using a Real-Time In Vivo Portal Dosimetry System. Technol Cancer Res Treat 2019;18:1-6.
18. Lim SB, Lee N, Zakeri K, Greer P, Fuangrod T, Coffman F, et al. Can the Risk of Dysphagia in Head and Neck Radiation Therapy Be Predicted by an Automated Transit Fluence Monitoring Process During Treatment? A First Comparative Study of Patient Reported Quality of Life and the Fluence-Based Decision Support Metric. Technol Cancer Res Treat 2021; 20: 1-7.
19. Castelli J, Simon A, Lafond C, Perichon N, Rigaud B, Chajon E, et al. Adaptive radiotherapy for head and neck cancer. Acta Oncologica 2018;57:1284–92.
20. Dewan A, Sharma S, Dewan A, Srivastava H, Rawat S, Kakria A, et al. Impact of Adaptive Radiotherapy on Locally Advanced Head and Neck Cancer - A Dosimetric and Volumetric Study. Asian Pac J Cancer Prev 2016;17:985–92.
21. Brown E, Owen R, Harden F, Mengersen K, Oestreich K, Houghton W, et al. Predicting the need for adaptive radiotherapy in head and neck cancer Radiother Oncol 2015;116:57–63.

22. Brown E, Owen R, Harden F, Mengersen K, Oestreich K, Houghton W, et al. Head and neck adaptive radiotherapy: Predicting the time to replan. Asia Pac J Clin Oncol 2016;12:460–7.
23. Tanooka M, Doi H, Ishida T, Kitajima K, Wakayama T, Sakai T, et al. Usability of Deformable Image Registration for Adaptive Radiotherapy in Head and Neck Cancer and an Automatic Prediction of Replanning. IJMPCERO 2017;06:10–20.
24. Morgan HE, Sher DJ. Adaptive radiotherapy for head and neck cancer. Cancers Head Neck 2020;5:1-16.
25. Zhang P, Simon A, Rigaud B, Castelli J, Ospina Arango JD, Nassef M, et al. Optimal adaptive IMRT strategy to spare the parotid glands in oropharyngeal cancer. Radiother Oncol 2016;120:41–7.
26. Kearney M, Coffey M, Leong A. A review of Image Guided Radiation Therapy in head and neck cancer from 2009–2019 – Best Practice Recommendations for RTTs in the Clinic. Tech Innov Patient Support Radiat Oncol 2020;14:43–50.
27. Cheng CS, Jong WL, Ung NM, Wong JHD. Evaluation of Imaging Dose From Different Image Guided Systems During Head and Neck Radiotherapy: A Phantom Study. Radiat Prot Dosimetry 2017;175:357–62.
28. Jarrett D, Stride E, Vallis K, Gooding MJ. Applications and limitations of machine learning in radiation oncology. Br J Radiol 2019;92:1-12.
29. Barragán-Montero A, Javaid U, Valdés G, Nguyen D, Desbordes P, Macq B, et al. Artificial intelligence and machine learning for medical imaging: A technology review. Phys Med 2021;83:242–56.
30. Deig CR, Kanwar A, Thompson RF. Artificial Intelligence in Radiation Oncology. Hematol Oncol Clin North Am 2019; 33:1095–104.
31. Sheng K. Artificial intelligence in radiotherapy: a technological review. Front Med 2020;14:431–49.
32. Siddique S, Chow JCL. Artificial intelligence in radiotherapy. Rep Pract Oncol Radiother 2020;25:656–66.
33. Giraud P, Giraud P, Gasnier A, El Ayachy R, Kreps S, Foy J-P, et al. Radiomics and Machine Learning for Radiotherapy in Head and Neck Cancers. Front Oncol 2019;9:174-186.
34. Marzi S, Pinnarò P, D'Alessio D, Strigari L, Bruzzaniti V, Giordano C, et al. Anatomical and Dose Changes of Gross Tumour Volume and Parotid Glands for Head and Neck Cancer Patients during Intensity-modulated Radiotherapy: Effect on the Probability of Xerostomia Incidence. Clin Oncol (R Coll Radiol) 2012;24:e54-62.
35. Langius JAE, Twisk J, Kampman M, Doornaert P, Kramer MHH, Weijs PJM, et al. Prediction model to predict critical

weight loss in patients with head and neck cancer during (chemo)radiotherapy. Oral Oncol 2016;52:91–96.

36. Pandit P, Patil R, Palwe V, Yasam VR, Nagarkar R. Predictors of Weight Loss in Patients With Head and Neck Cancer Receiving Radiation or Concurrent Chemoradiation Treated at a Tertiary Cancer Center. Nutr Clin Pract 2020; 35:1047–52.
37. Lee D, Zhang P, Nadeem S, Alam S, Jiang J, Caringi A, et al. Predictive Dose Accumulation for HN Adaptive Radiotherapy. Phys Med Biol 2020;65:1-27.
38. Pukala J, Johnson PB, Shah AP, Langen KM, Bova FJ, Staton RJ, et al. Benchmarking of five commercial deformable image registration algorithms for head and neck patients. J Appl Clin Med Phys 2016;17:25–40.
39. Hvid CA, Elstrøm UV, Jensen K, Alber M, Grau C. Accuracy of software-assisted contour propagation from planning CT to cone beam CT in head and neck radiotherapy. Acta Oncologica 2016; 55:1324–30.
40. La Macchia M, Fellin F, Amichetti M, Cianchetti M, Gianolini S, Paola V, et al. Systematic evaluation of three different commercial software solutions for automatic segmentation for adaptive therapy in head-and-neck, prostate and pleural cancer. Radiat Oncol 2012;7: 160-175.
41. Raju MP, Laxmi AJ. IOT based Online Load Forecasting using Machine Learning Algorithms. Procedia Comput Sci 2020; 171:551–60.
42. Kapoor NR, Kumar A, Kumar A, Kumar A, Mohammed MA, Kumar K, et al. Machine Learning-Based CO2 Prediction for Office Room: A Pilot Study. Wirel Commun Mob Comput 2022;2022: 1-16.
43. Barker JL, Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, et al. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol Biol Phys 2004;59: 960–70.
44. Bak B, Skrobala A, Adamska A, Kazmierska J, Jozefacka N, Piotrowski T, et al. Criteria for Verification and Replanning Based on the Adaptive Radiotherapy Protocol "Best for Adaptive Radiotherapy" in Head and Neck Cancer. Life (Basel) 2022;12:722-735.
45. Bhide SA, Davies M, Burke K, McNair HA, Hansen V, Barbachano Y, et al. Weekly Volume and Dosimetric Changes During Chemoradiotherapy With Intensity -Modulated Radiation Therapy for Head and Neck Cancer: A Prospective Observational Study. Int J Radiat Oncol Biol Phys 2010;76:1360–8.

46. Yao W-R, Xu S-P, Liu B, Cao X-T, Ren G, Du L, et al. Replanning Criteria and Timing Definition for Parotid Protection-Based Adaptive Radiation Therapy in Nasopharyngeal Carcinoma. Biomed Res Int 2015;2015:1-9.
47. Li J, Xia T, Yang X, Dong X, Liang J, Zhong N, et al. Malignant solitary pulmonary nodules: assessment of mass growth rate and doubling time at follow-up CT. J Thorac Dis 2018;10:S797–806.
48. Coca-Pelaz A, Takes RP, Hutcheson K, Saba NF, Haigentz M, Bradford CR, et al. Head and Neck Cancer: A Review of the Impact of Treatment Delay on Outcome. Adv Ther 2018;35:153–60.