Received: 2 April 2025 | Revised: 10 September 2025 | Accepted: 22 December 2025

DOI: 10.1002/acm2.70465

JOURNAL OF APPLIED CLINICAL MEDICAL PHYSICS

RESEARCH ARTICLE

Quantification of head and neck cancer patients' anatomical changes during radiotherapy: Toward the prediction of replanning need

Odette Rios-Ibacache[1] | **James Manalad**[1] | **Kayla O'Sullivan-Steben**[1] | **Emily Poon**[2] | **Luc Galarneau**[2] | **Julia Khriguian**[3] | **George Shenouda**[4] | **John Kildea**[1,2]

[1]Medical Physics Unit, McGill University, Montreal, QC, Canada

[2]Research Institute of the McGill University Health Centre, McGill University, Montreal, QC, Canada

[3]Department of Radiation Oncology, Hôpital Maisonneuve-Rosemont, Montreal, QC, Canada

[4]Department of Radiation Oncology, McGill University, Montreal, QC, Canada

Correspondence
Odette Rios-Ibacache, Medical Physics Unit, McGill University, Montreal, QC, Canada.
Email: odette.riosibacache@mail.mcgill.ca

Funding information
Réseau de cancérologie Rossy; Natural Sciences and Engineering Research Council of Canada; Institut de recherche, Centre universitaire de santé McGill; Fonds de Recherche du Québec–Santé

Abstract

Background: Head and neck cancer (HNC) patients undergoing radiotherapy (RT) may experience anatomical changes during treatment, which can compromise the validity of the initial treatment plan, necessitating replanning. However, ad hoc replanning disrupts clinical workflows and increases workload. Currently, no standardized method exists to quantify anatomical variation that necessitates replanning.

Purpose: This project aimed to create geometrical metrics to describe anatomical changes in HNC patients during RT. The usefulness of these metrics was evaluated by a univariate analysis and through machine learning (ML) models to predict the need for replanning.

Methods: A cohort of 150 HNC patients treated at McGill University Health Centre was analyzed. Based on the shapes of the RT structures (body, PTV, mandible, neck, and submandibular contours), we developed 43 metrics and automatically calculated them through a Python pipeline that we called *HNGeoNatomyX*. Univariate analysis using linear regression was conducted to obtain the rate of change of each metric. We also obtained the relative variation of each metric between the pre-treatment and replanning-requested scans. Fraction-specific ML models (incorporating information available up to and including the specific fraction) for fractions 5, 10, and 15 were built using metrics, clinical data, and feature selection techniques. Model performance was estimated with repeated stratified 5-fold cross-validation resampling technique and the area under the curve (AUC) of the receiver operating characteristic (ROC) curve.

Results: Univariate analysis showed that body- and neck-related metrics were most predictive of replanning need. Our best specific multivariate models for fractions 5, 10, and 15 yielded testing scores of 0.82, 0.70, and 0.79, respectively. Our models early predicted replanning for 76% of the true positives.

Conclusions: The created metrics have the potential to characterize and distinguish which patients will necessitate RT replanning. They show promise in guiding clinicians to evaluate RT replanning for HNC patients and streamline workflows.

KEYWORDS
geometrical metrics, head and neck cancer, machine learning, mass loss, radiotherapy, treatment replanning

1 | INTRODUCTION

Worldwide, more than 830 000 people are diagnosed with Head and Neck Cancer (HNC), and more than 430000 patients die from it each year.[1] Statistically, around 75% of patients with HNC have been reported to benefit from RT treatment as a primary or adjuvant treatment option.[2] However, patients with HNC who undergo RT may experience acute or late radiation toxicities during or after treatment, respectively, including dermatitis, xerostomia, and mucositis.[3,4] These toxicities may be painful and contribute to local infections and affect nutrition, as patients may have difficulty eating, drinking, and swallowing.[5] Consequently, weight loss and anatomical changes (volumetric or spatial) may occur. Furthermore, the tumor itself may respond well to radiation and shrink in size.

Anatomical changes during HNC RT are problematic, as they can compromise the delivery of the prescribed treatment dose (i.e., underdosing the tumor volume or overdosing the organs-at-risk (OARs)), necessitating treatment replanning.[6,7] To account for these changes, HNC patients are usually treated with an offline adaptive RT (ART) approach, using image-guided RT (IGRT) guidelines, which initially involves acquiring a CT simulation (CT sim) image using an immobilization device prior to the treatment and defining and outlining the planning target volume (PTV) and the OARs. Subsequently, during treatment, for each RT fraction, a pre-treatment on-board cone beam CT (CBCT) image is typically taken to confirm patient positioning and anatomy.

An example of the complications that HNC patients may experience is that, due to the weight loss, the PTV contour (considered initially) can be invalidated. Additionally, the thermoplastic treatment mask to immobilize can become loose. The mask looseness can, in turn, lead to unwanted movement and positioning difficulties, causing setup errors during treatment, which can then affect the delivered doses to many critical structures, such as the brainstem and spinal cord.[8] These types of changes can be seen in Figure 1, where a patient scan is presented over treatment fractions.

At our clinic, roughly 37.8% of HNC patients undergo RT replanning due to anatomical or dosimetric changes. According to the literature,[9–13] 4.5% - 49.3% of HNC patients require a replan. Figen et al. (2020)[9] reported that the most common replanning requests in the clinic are due to tumor shrinkage and weight loss, accounting for 71% of the requests.

Despite the widespread need to replan, there is no formal consensus or standardized guidance on when to replan in HNC RT. For example, Bhide et al. (2010)[14] noted that the parotid glands received an increased dose by the 4th week of treatment due to their medial shift, indicating that patients could benefit from replanning before week four. In contrast, Fiorentino A. et al. (2012)[15] suggested that a replan should be indicated by the 3rd week of treatment to avoid parotid gland overdose. Wang et al. (2010)[16] observed that, for patients with nasopharyngeal HNC, replanning before the 25th fraction of RT ensures safe doses to OARs. Furthermore, no standard method has been established to define a threshold amount of mass loss or anatomical change in the head and neck region that triggers a replan.[17]

Additionally, at the clinic, the planning workflow is typically complex, and various workloads and economic factors[18] can cause the replanning process to be delayed until absolutely necessary.[19] Clinical personnel, including radiation therapists and medical physicists, must work together to schedule and perform CT rescanning, recontouring, and replanning effectively.[20]

To address this issue, some studies, such as Brown et al. (2016),[21] have tried to predict the *optimal time* for replanning by attempting to identify *pre-treatment factors* that influence the replanning decision. Their results showed that nasopharyngeal cases require replanning earlier than oropharyngeal cases (3rd and 4th weeks of treatment, respectively). In the context of the possible use of artificial intelligence (AI) methods to predict RT replanning for HNC, a recent study by Chinnery et al. (2024)[10] showed that radiomic and dosimetric information from initial CT sim scans has the potential to aid replanning predictions. However, their model did not incorporate information on the effect of IMRT on the patient and did not consider subsequent medical images after the initiation of treatment.

The lack of a tool that can systematically record and characterize the anatomical and dosimetric changes that HNC patients experience during RT makes it difficult to assess which patients may benefit from replanning and when. We define the need for replanning as the time point when image-guided positioning can no longer ensure adequate coverage of the PTV during treatment. With this in mind, our study aimed to create metrics that quantify anatomical changes based on the geometry of the HN region and examine their predictive values, in combination with clinical data, for RT replanning in HNC patients using a machine learning (ML) model. We surmised that the development of such metrics and the creation of an automatic pipeline to calculate and evaluate them could serve as a supportive decision

JOURNAL OF APPLIED CLINICAL MEDICAL PHYSICS

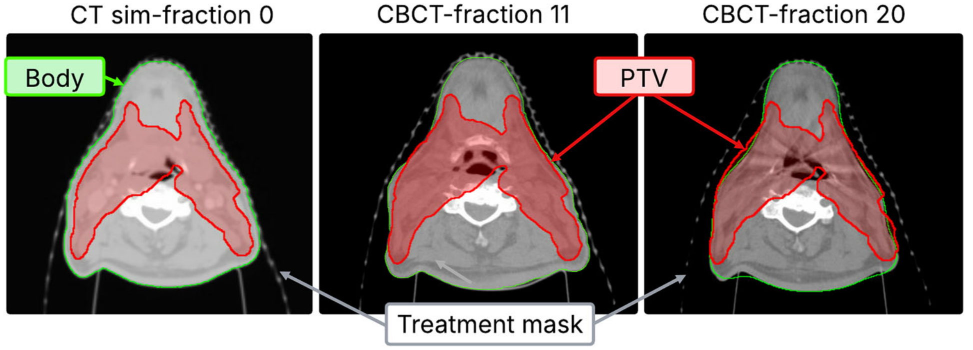

FIGURE 1 HNC patient's axial CT sim and CBCT scan slice. The patient was replanned in fraction 20 out of 35 at our institution. Body and PTV contours over the RT treatment. The mass loss, the mask looseness, and the invalidated PTV contour can be seen. HNC, head and neck cancer; CBCT, cone beam CT; PTV, planning target volume; RT, radiotherapy.

aid for clinicians over the course of the treatment of HNC patients.

2 | MATERIALS AND METHODS

2.1 | Patient cohort and study sample

This retrospective study considered patients diagnosed with HNC or unknown primary tumors in the head and neck region. These patients were diagnosed and biopsy-proven with squamous cell carcinoma (SCC) or adenocarcinoma. Data were collected for a cohort of 381 patients who received or started receiving intensity modulated RT (IMRT) treatment at the MUHC between January 1, 2017 and March 31, 2024. Treatment regimens included RT alone or in combination with systemic therapies, such as chemotherapy and targeted therapy. The prescribed RT doses in our patient cohort consisted of 70 Gy in 35 fractions, 60 Gy in 30 fractions, and 66 Gy in 33 fractions to the PTV, with curative intent. This retrospective study was approved by the Research Ethics Board (REB) of McGill University Health Centre. All of the study's work was conducted in accordance with REB guidelines.

In our institution, the reason for HNC RT replanning is documented in the electronic medical record, which includes the selection of checkboxes (change in target volume, weight loss, change in skin separation, alteration in muscle mass and fat, fluid shifts, change in immobilization device, difficulties in setup, and other reason), and free text to add more details. The final decision is then usually discussed, and the treatment proceeds as decided (to replan or not). We retrospectively analyzed the reasons provided by either the radiation oncologist, the radiation technologist, or the medical physicists to replan. Patients' inclusion and exclusion criteria were established based on the number of CBCT images acquired, the reported reasons for replanning, and whether or not patients finished their treatments.

Since anatomical changes have been reported to occur during the first week of treatment,[22] and our institutional imaging guidelines specify that at least two CBCTs should be taken per week for IMRT, we defined an inclusion criterion that each patient was required to have a CT sim image and at least two CBCT images taken on two different fractions during treatment. The replan request fraction had to be within the fraction range of the first and last CBCT images or correspond to one of them. Furthermore, we excluded cases where the reported replanning reasons were not related to the patient's anatomy. These reasons include holes that developed in the treatment masks, patients who did not tolerate the treatment mask, or patients who were unable to tolerate treatment for psychological reasons. The dataset selection process is summarized in Figure S-1.

Among the 362 eligible patients, data collected comprising 8 years, 37.8% ($n = 137$) had their RT treatments replanned, and 62.2% ($n = 225$) completed their RT course without replanning. Due to time constraints, a sample of 75 replanned cases and 75 non-replanned cases was randomly selected from the eligible patient cohort to provide a balanced class distribution for the ML analysis.[23]

2.2 | Data collection

CT sim images, taken at the beginning of treatment planning, and CBCT images, taken over the course of treatment, were collected from the Treatment Planning System (TPS), Eclipse v.15 (Varian Medical Systems, Inc., Palo Alto, CA) and saved in our institution's server. For each patient, all the available medical images for the treatment course were exported from the TPS as DICOM files, and all RT data were exported as

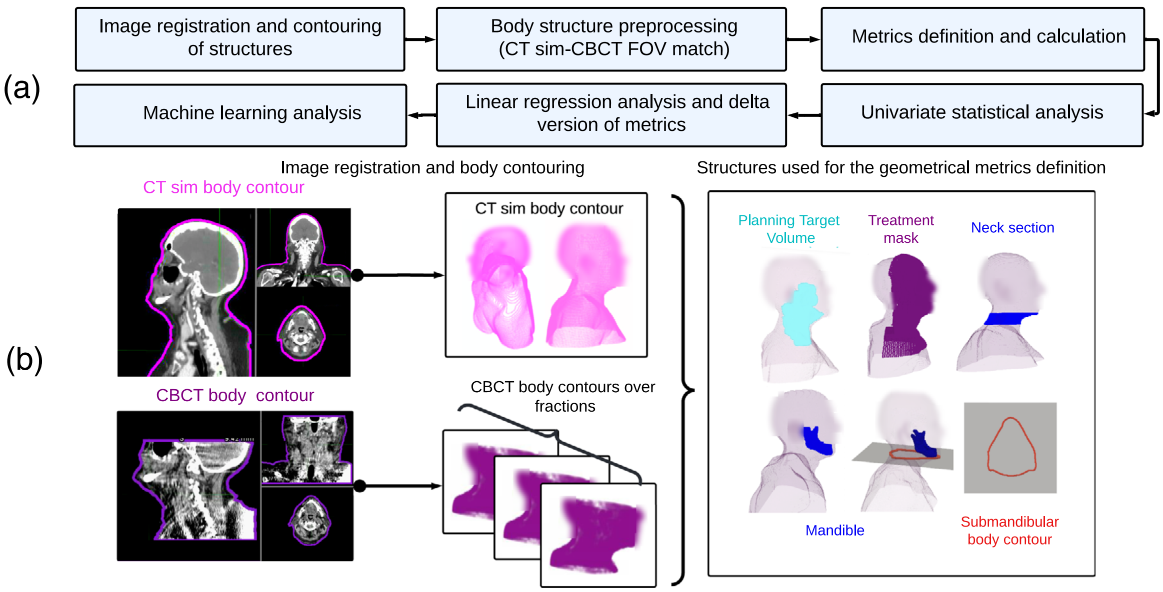

FIGURE 2 (a) shows the workflow describing the steps that were followed for the metrics definition and statistical analysis, beginning with the image registration process and finishing with the machine learning analysis. (b) shows the generated 3D structures used for our analysis, including the CT sim and CBCTs body contouring, and geometrical metrics definition.CBCT, cone beam CT; CT sim, CT simulation.

DICOM-RT files. All data were anonymized during the TPS export.

In our cancer centre, clinicians use questionnaires, recorded at regular intervals throughout RT treatment, to collect certain temporally variable clinical information for HNC patients. These data include weight, toxicity grades (e.g., dysphagia, mucositis, xerostomia, dermatitis, and laryngeal side effects), Karnofsky performance status (KPS), hematological test results, and presence or absence of tube feeding (percutaneous endoscopic gastrostomy (PEG)). These data were included to allow a thorough evaluation of the clinical evolution of each patient over time.

This study also incorporated time-independent clinical information about the patient, including the status of the tumor suppressor gene p16 (a surrogate marker for HPV status[24]), smoking or tobacco use history, TNM staging, overall stage of cancer, and age at the start of treatment. We included these time-independent variables since previous research has found HPV-positive patients with HNC experience significantly more malnutrition and weight loss when they go under RT treatment.[25,26] Likewise for cases of smoking history and radiation-induced dysgeusia.[27]

Three-dimensional geometrical information was analyzed using medical images, RT structures, and patient body shapes to study the anatomical changes that occurred during RT treatment (Figure 1). The following subsections describe the process of contouring the external body and the treatment mask, which were then used to create metrics to characterize anatomical changes. Figure 2 summarizes the workflow of this study.

2.3 | Image registration and contouring of structures

For each patient, we registered each of their CBCT images to their CT sim DICOM image to align with the RT structures corresponding to the PTV and OARs, using a rigid registration method. For the replanned patients, only their pre-replanning CBCTs were registered. The body contour of each CBCT image was generated using the commercial contouring application MIM MAESTRO v7.1.6 (MIM Software Inc., Cleveland, OH). To ensure that the body structure was correctly contoured, each slice was visually examined. Slices in which the patient's imaged anatomy was partially imaged were not included in the body contour. The final contoured image was saved in the DICOM-RT structure set format. Since the goal of this study was to predict replanning, and considering that replanning often occurs before the 4th week of treatment,[14,28] the registration process was only performed for fractions before or including the 25th fraction.

Although treatment mask looseness is sometimes the reason for replanning,[9] the mask is not included as an RT structure in clinical practice. To address this, a semi-automatic 2D contouring algorithm was developed in Python v3.7 to define the mask structure. The algorithm receives as input the array of the CT sim image slices, the index number of the slice to be contoured, the CT sim body contour (used to calculate the superior and inferior limits), and a threshold value. The output points are saved in the same coordinate system as the RT structure set of each patient. The resulting matrix of points

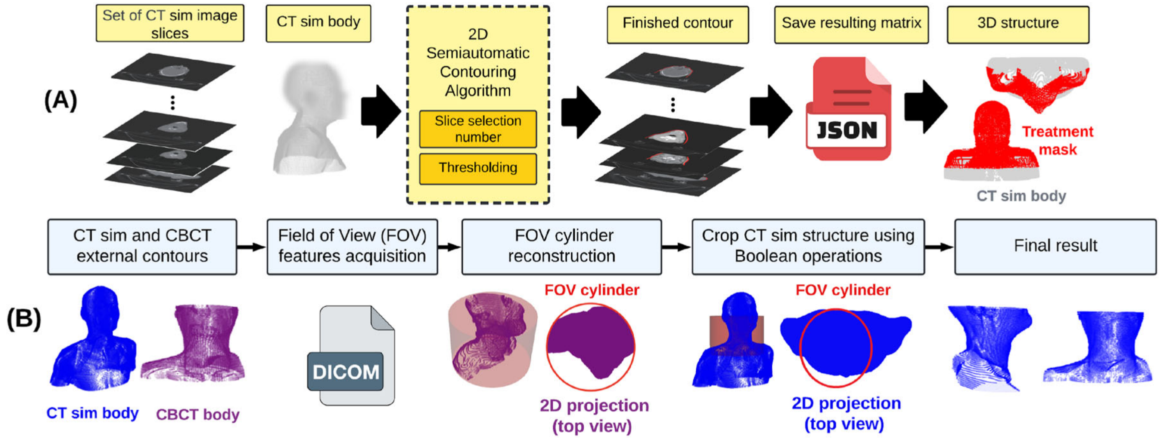

FIGURE 3 (a) Flowchart of the treatment mask contouring steps. (b) A diagram describing the CBCT image FOV reconstruction procedure and the CT sim contour cropping procedure to match the geometrical shape.CBCT, cone beam CT; FOV, field of view.

is stored in the *JSON* file format. The summarized steps can be seen in Figure 3a.

2.3.1 | Body structure preprocessing

Since the cone-shaped beam used to acquire a CBCT image limits its field of view (FOV), portions of a patient's anatomy may be missing from the CBCT when compared to the corresponding CT sim image, which has a much wider FOV. Thus, to ensure consistent anatomy across fractions for a given patient, a preprocessing step was applied to reconstruct the CBCT FOV and use it to crop the CT sim's contour.

The acquisition isocenter of the first CBCT (fraction 1) was extracted from its DICOM-RT structure Set file (as previously registered to the CT sim using MIM MAESTRO). The reconstruction diameter stored in the CBCT DICOM information was retrieved using the open-source library *pydicom*[29] to define the outer boundary and reconstruct the FOV. Next, an algorithm based on Boolean operations was created to crop the CT sim's body contour for each patient. The process can be seen in Figure 3b. Additionally, since each CBCT image captures the anatomy for slightly different *z*-ranges (height), each body contour was trimmed in the *z* direction to the common overlapping *z*-region across all CBCTs for a given patient.

2.4 | Metrics to describe patient anatomy

Based on the 3D shape and the 2D contours of the body and RT structures (including the PTV and mandible), we defined a total of 43 quantitative and continuous metrics to describe the anatomical changes of the patient during RT delivery. These changes include those observed from CT sim to CBCT, as well as between successive CBCTs (see Figure 1). We grouped the metrics into six categories, which are summarized in Table 1 and detailed in the following subsections.

2.4.1 | Body-related metrics

This set of six metrics was based on the body region covered and the differences between each CBCT body contour and the corresponding CT sim body contour. These metrics are labeled with the subscript *Body* in Table 1 and are visually shown in Figure S-2.

We defined a volume metric (V_{Body}) to quantify the volume enclosed by the body contour. This metric was calculated on the basis of the total number of voxels enclosed by the body/external contour multiplied by the voxel volume (based on the CT sim image resolution: pixel spacing and slice thickness).

Next, a series of 3D and 2D distances was established to measure the change for each fraction relative to the CT sim. These metrics were inspired by Bivrio D. et al.(2018),[30] who calculated the differences between body contours in cylindrical coordinates, and the analysis that physicists and radiation oncologists perform during clinical evaluation (a slice-by-slice analysis of images). For the 3D distance CD_{Body}, we used the *chamfer distance* module from the open-source library *point-cloud-utils*.[31] The 3D maximum distance was calculated by obtaining the maximum value of the distances between the nearest neighbors of the points in the contour using *KDtree* from *scipy*.[32]

To calculate 2D distances, each axial slice of the contour was analyzed individually. In each slice, the distance between the points was calculated using the *directed hausdorff* module from *scipy*. The maximum, mean, and median distances for each CBCT were

TABLE 1 Metric definitions that were created to analyze the anatomical changes of replanned and non-replanned patients.

Metric	Symbol	Metric	Symbol
(1) Body-related metrics		Avg. distance from the mandible to body	M_{avg}
Body volume	V_{Body}	SD of distances from the mandible to body	σ_M
Chamfer or avg. distance (3D)	CD_{Body}		
Haussdorff or max. distance (3D)	HD_{Body}	*(5) Neck-related metrics*	
Max. 2D distance	D_{Body}	Neck volume	V_{Neck}
Median 2D distance	$\tilde{D}_{Body}$	Chamfer or avg. neck distance	CD_{Neck}
Avg. 2D distance	$\bar{D}_{Body}$	Haussdorff or max. neck distance	HD_{Neck}
(2) Treatment mask-related metrics		Max. 2D neck distance	D_{Neck}
Max. distance to treatment mask	$max\{B_{mask}\}$	Median 2D neck distance	$\tilde{D}_{Neck}$
Avg. distance to treatment mask	$\bar{B}_{mask}$	Avg. 2D neck distance	$\bar{D}_{Neck}$
SD of distances to treatment mask	$\sigma_{B_{mask}}$	Min. 3D neck radius	R^{3D}_{min}
Air volume between body and treatment mask	$V^{air}_{Body-to-mask}$	Max. 3D neck radius	R^{3D}_{max}
(3) PTV-related metrics		Avg. 3D neck radius	R^{3D}_{avg}
Min. distance from CT sim PTV to CBCT body	x^{PTV}_{min}	Ratio between min. and max. 3D neck radius	φ^{3D}_R
Max. distance from CT sim PTV to CBCT body	x^{PTV}_{max}	Avg. cross-sectional neck area	A^{2D}_{avg}
Avg. distance from PTV to body	x^{PTV}_{avg}	Neck surface area	SA_{Neck}
Median distance from PTV to body	x^{PTV}_{med}	Neck compactness	C_{Neck}
SD of the distances from PTV to body	x^{PTV}_{std}	*(6) Submandibular-related metrics*	
Volume PTV inner	VI_{PTV}	Submandibular area	A_{sub}
Volume PTV outer	VO_{PTV}	Min. 2D submandibular radius	R^{sub}_{min}
Volume PTV inner ratio	$VI_{PTV} : LV_{Body}$	Max. 2D submandibular radius	R^{sub}_{max}
Volume PTV outer ratio	$VO_{PTV} : LV_{Body}$	Avg. 2D submandibular radius	R^{sub}_{avg}
(4) Mandible-related metrics		Ratio between R^{sub}_{min} and R^{sub}_{max}	φ^{2D}_{Rsub}
Min. distance from the mandible to body	M_{min}	Maximum longitudinal chord	l^{sub}_y
Median distance from the mandible to body	M_{med}	Maximum lateral chord	l^{sub}_x

Abbreviations: PTV, planning target volume; SD, standard deviation.

determined from the distribution of distances across all slices for that CBCT.

2.4.2 | Treatment mask-related metrics

A 3D structure of the treatment mask was generated using *pyvista*,[33] based on its contour. Next, to track the looseness of the mask, we created metrics to relate the treatment mask to the body contour. Four metrics were defined for this purpose: three based on distances that were calculated using the *KDTree* and one volume that describes the space between the body and the mask, which was calculated using the same approach as for V_{Body}. The metrics were labeled with the subscript *mask* and are visually represented in Figure S-3.

2.4.3 | PTV-related metrics

Motivated by Bartosz B. et al. (2022),[34] who found that the extension of the PTV outside the body on CBCTs is the reason for a high number of replanning decisions, we defined metrics to describe the location of the PTV (originally generated on the CT sim image) relative to the body (Figure 1). These served as indicators of potential inadequate target coverage during treatment. Five metrics were created to describe the distance between the PTV and the body contour, as well as the portions of the PTV inside and outside the body.

To determine the portion of the PTV inside and outside the body, the *compute implicit distance* and *threshold* modules from the *pyvista* library were used. These functions divide a 3D structure according to a surface, allowing the identification of inner and outer PTV points to calculate the distances. The PTV-related metrics were labeled with the subscript *PTV* in Table 1. The defined PTV-related metrics are visually described in Figure S-4.

2.4.4 | Mandible-related metrics

Another way to characterize the mass loss that patients with HNC may experience is to track the distance between the body contour and a rigid bone structure that experiences minor dimensional changes. This type of bone structure should remain or be fixed in the same position within the body over the course of treatment. The mandible is one of these types of bone structure, and since it was included in the CT sim RT structure set file for all patients in our cohort, it served as a reliable reference point to track mass loss in the face region. For each patient, distances were calculated between the mandible and the body contour of the patient. The *compute implicit distance* and *threshold* modules from the *pyvista* library were again used for this purpose. The metrics were labeled with the letter and subscript M.

2.4.5 | Neck-related metrics

Inspired by studies that have shown that patients undergoing RT experience changes in the dimensions of their neck regions,[35,36] our study included the creation of an algorithm to determine consistent boundaries of the neck region, as it was not previously explicitly contoured. The procedure involved the identification of a z_b value from the most inferior slice of the body point cloud that does not include any shoulder anatomy and is at the boundary of the FOV cylinder. To consider a consistent neck region across the CBCTs of an individual patient, the region was defined at the superior end by the slice below the mandible and at the inferior end by the most inferior point determined from the CBCT body contours' z-values, z_b. To avoid the inclusion of the shoulders, a three-slice margin was employed above the most inferior point. Metrics were defined to describe the size and shape of the neck using similar calculation methods used for the body-related metrics. They are presented in Table 1 using the subscript *Neck*.

2.4.6 | Submandibular-related metrics

To develop comprehensive variables that describe changes in all possible directions and facilitate a thorough analysis, additional 2D metrics related to the submandibular region were created. The submandibular region is part of the neck and can experience a high rate of change during treatment. To calculate these metrics, the submandibular plane was determined along with the corresponding body contour points. To distinguish the metrics, the symbols were labeled with the superscript or subscript *sub*, as appropriate. In Figure S-5, the metrics can be visualized.

To systematically track changes over time using all defined metrics, an automatic extraction pipeline (referred to as *HNGeoNatomyX*) was developed using the Python v3.7 scripting language. This tool facilitated the management, extraction, and analysis of all the metrics described above.

2.5 | Univariate analysis

Using our automatic metrics extraction pipeline and a CPU corresponding to an Intel Core Processor (Skylake) with 8 GB of RAM, we analyzed the data of our 150 patients. A univariate statistical analysis was performed to provide information on the potential predictive power of each extracted metric and its ability to guide replanning. This type of analysis included a linear regression and a delta analysis. Linear regression was used to individually examine the evolution of each metric over time, while delta analysis was performed to quantify the anatomical change at the time of the replanning request.

2.5.1 | Linear regression analysis

This analysis involved the calculation of the rate of change of a given metric to see how it changed across fractions for each of the two classes of patients: replanned and non-replanned. Specifically, the rate of change was obtained from the slope calculation up to each fraction number. As reported in previous studies,[37–39] weight loss and decreased neck area over time were more prominent in replanned patients compared to non-replanned patients. Thus, we hypothesized that the rate at which the metrics change may provide useful information.

We generated graphs to visually inspect the longitudinal data across treatment fractions and show the metrics' progression over time. Quantitatively, each metric was analyzed using the non-parametric Mann–Whitney U (MWU) test to determine if the distributions corresponding to the classes replanned versus non-replanned were statistically different. A Bonferroni correction was applied to account for the application of multiple hypothesis tests.[40] Using the definition of the MWU test, the predictive power of each metric was calculated for the fractions for which the test had p-values below 0.05, and an average was obtained. The predictive power was calculated using the AUC value from the U–statistic for ROC curves.[41] A 95% confidence interval (C.I.) was established through bootstrapping with 100 repetitions. The same procedure was used to compare the results for the weight values (also collected over time).

2.5.2 | Delta analysis

Relative differences between the metric values at the time of the CT sim (fraction 0) and a fraction of interest were also considered in our analysis, since these values are frequently reported in literature.[37] These quantities were referred to as the Delta (Δ) of each metric and were defined as $\Delta p_{fx} = \frac{p_{fx} - p_0}{p_0}$. Where p_{fx} is the metric's value at the fraction of interest (fx) and p_0 is the

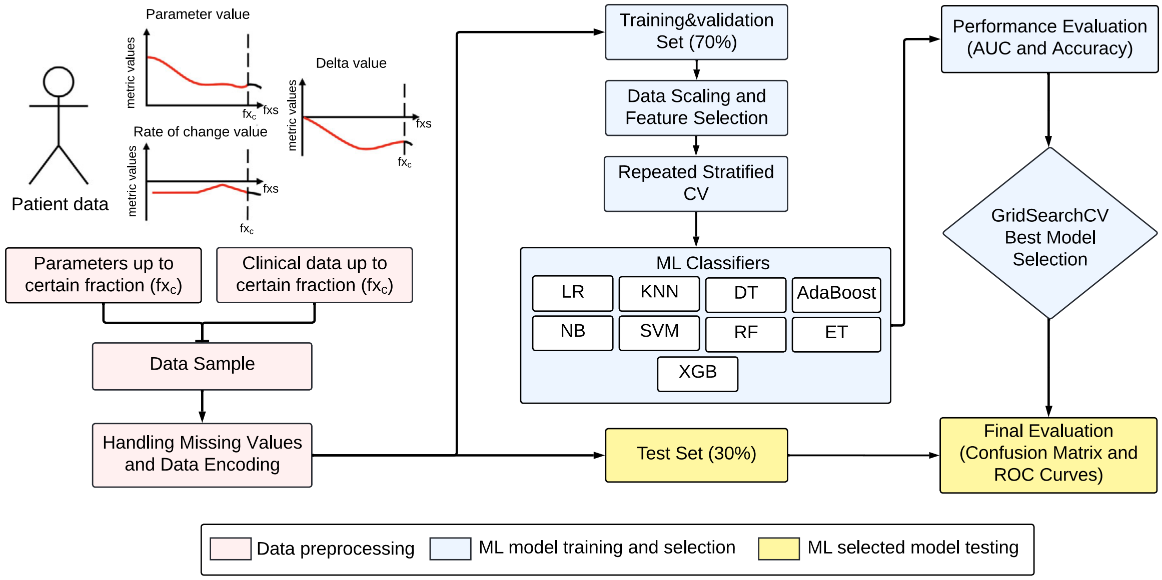

FIGURE 4 ML pipeline analysis conducted in our research. CV is cross-validation, and GridSearchCV is the hyperparameter tuning method. StandardScaler: Data scaling or standardization method used. LR: Logistic regression, KNN: KNearestNeighbors, DT: DecisionTree, NB: NaiveBayes, SVM: Support vector machine, RF: Random forest, ET: ExtraTree, XGB: eXtreme gradient boost classifier, fxs: fractions, fx_c certain fraction of interest.

metric's value at fraction 0 (CT sim). If the values of the metric at fraction 0 are zero, the Δ is instead calculated using the values of fraction 1. A similar approach has been employed in other studies, such as Delta radiomics for outcome prediction.[42,43] Histograms of the Δ values, representing the amount of anatomical change by the time of the replanning request, were generated to quantify the changes across the replanned patient sample.

2.6 | ML models to predict if replanning will be needed

After arranging and collecting the various metrics and clinical data, including the toxicity grades and patient weight at each fraction, missing values were identified and handled. Data were missing when the clinical questionnaires were not answered at every fraction or when not all patients had a CBCT taken at every fraction.

To handle missing values for individual patient paths (time series), three techniques were used: (1) filling the gap with the previous value in the series (in the case of toxicity grading), (2) filling the gap using linear interpolation (in the case of anatomical and geometrical metrics), and (3) filling the gap with the average of the data at a certain time point of the series. Lastly, categorical data, including TNM staging, sex, p16 status, smoking history, type of concomitant therapy, and HNC subsites, were handled using one-hot encoding. Ordinal encoding was used to preserve the hierarchy of the labels used for the overall cancer stage.

Figure 4 presents the pipeline used for the development and analysis of ML models. To build ML models as a prediction tool to identify which patients would likely need replanning, nine classifiers were evaluated: Logistic regression (LR), KNearestNeighbors (KNN), Decision Tree (DT), AdaBoost, Naive Bayes (NB), Support vector machine (SVM), Random forest (RF), ExtraTree (ET), and eXtreme gradient boost (XGB). Three specific fraction models were built using data up to and including fraction 5, fraction 10, and fraction 15, corresponding to the first, second, and third weeks of treatment. We selected these three timepoints since previous research found that patients experience anatomical variations from the first week that persist throughout treatment, with the most noticeable changes occurring mid-treatment.[22,44] Additionally, patients may benefit from replanning before the third week of treatment.[45]

To avoid overfitting, we split the data into 70%–30% training & validation and test (hold-out) sets. Only the information available up to and including the fraction of interest was considered. To simulate real-world situations where the RT fraction at which the patient is being treated is the one we would like to evaluate, we filtered the training & validation-testing split to include only the patients who had a request for replanning at or after the fraction for which the specific-fraction model is built.

To reduce the impact of data skewness or outliers and avoid data leakage,[46] the StandardScaler method from *scikit-learn* was used for data scaling. Feature selection (FS) was performed using both filter and wrapper methods, specifically SelectKBest and RFE, respectively. Since the RFE technique uses a classifier algorithm, in our case, we used RandomForest as suggested and used by Chen et al. (2018),[47] Darst et al. (2018),[48] and Wang et al. (2022).[49] All available data (metrics, rate of change associated-metrics, Δ associated-metrics, and clinical information) up to each fraction-specific time point were used as input for the FS techniques, allowing the algorithm to freely choose among the absolute number of features i. The i number was set to the number of metrics (43) plus the number of clinical characteristics and toxicities (21), for a total of $i = 64$ features. A subset comprising half of the total of inputs ($i = 32$) was also tested to explore options with fewer degrees of freedom. FS techniques were only applied to the training set.

A repeated stratified K-fold cross validation (CV)[50] method was used during the training & validation procedure to maintain the class proportion using $k = 5$ with 10 repetitions. The performance of the ML models was evaluated using accuracy and AUC score. In each CV loop, performance metrics were obtained, along with their average, standard deviation, and 95% confidence interval. A GridSearchCV hyperparameter tuning was done for the best-scoring models (based on the highest average AUC values) to improve performance. After model selection, the test set was used to estimate the final performance of the model, using the ROC curve, AUC, accuracy, and confusion matrix.

3 | Results

3.1 | Study sample

The study sample comprised various sites of diagnosis: oropharynx (44.7%, $N = 67$), oral cavity (52.0%, $N = 39$), larynx (28.0%, $N = 21$), nasopharynx (14.7%, $N = 11$), sinus (2.0%, $N = 3$), nasal cavity (1.3%, $N = 2$), hypopharynx (1.3 %, $N = 2$), and others (3.33%, $N = 4$), which included diagnoses related to the pharynx ($N = 1$), cervical lymph nodes ($N = 1$), and malignant neoplasm without specification of site in the head, face and neck region ($N = 2$). This dataset consisted of male (72.0%, $N = 108$) and female (28.0%, $N = 42$) patients who were treated with total prescribed doses of 70 Gy (83.33%, $N = 100$) in 35 fractions, 60 Gy (10.83%, $N = 13$) in 30 fractions, and 66 Gy (5.83%, $N = 7$) in 33 fractions to the PTV, with curative intent.

The treatment regimens included RT alone (34.0%, $N = 51$) as well as combined with systemic therapies, such as chemotherapy (62.0%, $N = 93$), targeted therapy (2.0%, $N = 3$), and immunotherapy (2.0%, $N = 3$). The T stage patient distribution corresponded to T4 (36.7%, $N = 55$), T3 (26.0%, $N = 39$), T2 (18.7%, $N = 28$), T1 (14.7%, $N = 22$), T0 (1.3%, $N = 2$), TX (2.0%, $N = 3$), and one unknown case (0.7%, $N = 1$). The N stage distribution was N3 (9.3%, $N = 14$), N2 (45.3%, $N = 68$), N1 (16%, $N = 24$), N0 (28.7%, $N = 43$), and unknown cases (0.7%, $N = 1$). The M stage distribution was M0 (92.0%, $N = 138$), M1 (4.7%, $N = 7$), MX (2.0%, $N = 3$), and unknown cases (1.3%, $N = 2$). The characteristics of the patient sample, separated into replanned and non-replanned patients, are presented in Table S-1.

In this sample, the average number of CBCT images taken per patient were: 15 for non-replanned patients and 8 for replanned patients. The majority of the reported reasons for replanning were weight loss (70.7%, $N = 53$), followed by the PTV being in the air (10.7%, $N = 8$), change in skin separation (13.3%, $N = 10$), skin reaction or swelling (10.7%, $N = 8$), mask looseness (10.7%, $N = 8$), and change in target volume (8.0%, $N = 6$). The most frequent fraction in which patients had a replanning request was 15, and the average fraction was 13. While the average number of delivered fractions was 18. Note that several patients had more than one reason for replanning recorded in their charts. The complete distribution of the number of CBCTs taken, replan request fractions, and the number of RT fractions delivered of the replanning cohort is shown in Figure S-6.

3.2 | Univariate analysis

Using our automatic extraction pipeline, all 43 metrics were calculated for the 150 patients in our study sample. This took an average of 25 min per patient using our computational setup. The results of the MWU test, as applied to the rate of change up to each fraction for each metric, are presented in Table 2 (fractions at which *p-value*<0.05). The average AUC values for the rate of change in the fractions that showed differences are also presented, along with the 95% C.I. Additionally, the amount of change in each metric (using the delta analysis approach) that replanned patients experienced at the fraction at which the replanning was requested is presented with the symbol Δ. The best univariate results for each category of metrics (based on the closest AUC values to 1) are graphically presented in Figure S-3. In the case of the patient's rate of weight loss, the difference between the two groups was not notable until fraction 18, with an average AUC of 0.67 [0.58-0.75]. See Figure S-7 i.

3.3 | Machine learning analysis

A total of 150 HNC patients were used for the ML analysis, corresponding to 75 replanned and 75 non-replanned patients. The split (70%/30%) of the data into training & validation, and test sets corresponded to 105 patients and 45 patients, respectively. 105 patients (mean age of 62.9 years, 72 males and 33 females, 76 smokers, and 43 p16 positive) were included for training and CV, and 45 patients (mean age 63.6 years, 36 males and 9 females, 21 smokers, and 22 p16 positive) were included in the test set (see Tables S-2 and S-3).

TABLE 2 Univariate analysis results for each created geometrical metric.

Metric	slopes fx	avg. AUC	Δ	Metric	slopes fx	avg. AUC	Δ
(1) Body-related metrics				M_{avg}	7 → 25	0.72 [0.63-0.79]	−4.7%
V_{Body}	6 → 25	0.75 [0.69-0.83]	−3.9%	σ_M	ns	—	0.8%
CD_{Body}	3, 7 → 25	0.77 [0.72-0.84]	38.7%				
HD_{Body}	12 → 25	0.73 [0.64-0.80]	12.4%	*(5) Neck-related metrics*			
D_{Body}	9, 13 → 25	0.75 [0.70-0.81]	14.4%	V_{Neck}	6 → 25	0.74 [0.66-0.80]	−6.1%
$\tilde{D}_{Body}$	9 → 25	0.80 [0.74-0.85]	17.9%	CD_{Neck}	6 → 25	0.78 [0.72-0.83]	62.8%
$\bar{D}_{Body}$	7 → 25	0.80 [0.74-0.85]	16.8%	HD_{Neck}	21 → 25	0.71 [0.63-0.78]	22.6%
(2) Treatment mask-related metrics				D_{Neck}	15, 18 → 25	0.69 [0.61-0.77]	22.3%
$max\{B_{mask}\}$	ns	—	7.9%	$\tilde{D}_{Neck}$	14 → 25	0.76 [0.69-0.83]	34.4%
$\bar{B}_{mask}$	4 → 25	0.74 [0.66-0.81]	30.2%	$\bar{D}_{Neck}$	14 → 25	0.77 [0.69-0.84]	30.9%
$\sigma_{B_{mask}}$	ns	—	5.2%	R^{3D}_{min}	8 → 25	0.69 [0.63-0.78]	−6.6%
$V^{air}_{Body-to-mask}$	6 → 25	0.72 [0.64-0.79]	26.5%	R^{3D}_{max}	19 → 25	0.68 [0.61-0.76]	−1.0%
(3) PTV-related metrics				R^{3D}_{avg}	4, 6 → 25	0.78 [0.75-0.84]	−4.3%
x^{PTV}_{min}	6 → 25	0.77 [0.70-0.82]	−204.1%	φ^{3D}_{R}	ns	—	6.3%
x^{PTV}_{max}	ns	—	1.0%	A^{2D}_{avg}	6 → 25	0.75 [0.69-0.82]	−5.4%
x^{PTV}_{avg}	4, 6 → 25	0.76 [0.70-0.84]	−12.5%	SA_{Neck}	6 → 25	0.69 [0.64-0.75]	−0.7%
x^{PTV}_{med}	6 → 25	0.76 [0.68-0.83]	−15.5%	C_{Neck}	ns	—	13.2%
x^{PTV}_{std}	ns	—	2.2%	*(6) Submandibular-related metrics*			
VI_{PTV}	ns	—	−8.0%	A_{sub}	3 → 25	0.75 [0.70-0.81]	−5.2%
VO_{PTV}	4 → 25	0.70 [0.63-0.77]	—	R^{sub}_{min}	6 → 8, 10 → 25	0.68 [0.61-0.77]	−6.0%
$VI_{PTV} : LV_{Body}$	1, 3, 8 → 25	0.79 [0.73-0.85]	326.0%	R^{sub}_{max}	16 → 25	0.67 [0.59-0.75]	−0.6%
$VO_{PTV} : LV_{Body}$	4 → 25	0.69 [0.62-0.76]	—	R^{sub}_{avg}	6 → 25	0.75 [0.68-0.81]	−2.6%
(4) Mandible-related metrics				$\varphi^{2D}_{R_{sub}}$	18 → 19	0.65 [0.58-0.73]	6.4%
M_{min}	9 → 25	0.67 [0.60-0.75]	−24.8%	I^{sub}_{y}	6, 8 → 25	0.70 [0.63-0.77]	−1.6%
M_{med}	10 → 25	0.71 [0.61-0.79]	−3.6%	I^{sub}_{x}	6 → 10, 13 → 25	0.68 [0.61-0.75]	−2.8%

Where the ‘slopes fx’ (with fx: fraction) columns indicate the fractions over which the results of the MWU tests applied to the slope calculations of the metrics were statistically significantly different (*p-value*<0.05) for the replanned and non-replanned populations. ns: no significant results (*p-value*>0.05). The → symbol indicates a range of fractions from a starting point to an ending point. The ‘avg. AUC’ is the average AUC for these selected fractions AUC, along with its 95% confidence interval.

Models were built for three specific fractions (5, 10, and 15) to see if the metrics have the potential to predict the need for replanning using all the available data up to and including (but not after) the specified fraction. The models were tested using $n = 40$, $n = 36$, and n = 32 patients, respectively. The results are presented in Figure 5. Fraction 15-ML showed the lowest false positive and false negative cases. For more details about each model (including selected geometrical metrics by the FS techniques, the training performance of the models, and the selected hyperparameters for the best model after the GridSearchCV tuning process), see Tables S-4, S-5, S-6, and S-7 and Figures S-9, S-10, and S-11.

4 | Discussion

Several authors have attempted to identify metrics that may influence RT replanning decisions in HNC patients. They focused on information available before treatment, such as TNM stage, overall cancer stage, and demographic information. According to Nuyts et al. (2024), Hu et al. (2018), and Chen et al. (2014),[51–53] replanned patients typically had, at the beginning of treatment, larger values in body weight, tumor volumes, and parotid gland volumes. However, no definitive method has been established to determine metrics related to the amount of weight loss or anatomical change in the HN region that requires replanning. Therefore, a tool that facilitates the identification of patients who would likely benefit from replanning would be very useful in the clinic.

In this research study, we developed a series of 43 metrics related to patients’ 3D and 2D geometrical shapes to describe the anatomical variation experienced by HNC patients during RT. Our work included the creation of an extraction tool called *HNGeoNatomyX* that enables efficient assessment across fractions. The univariate and ML results presented demonstrate that these metrics can describe the anatomical evolution of replanned and non-replanned patients, showing their potential to differentiate between the two groups.

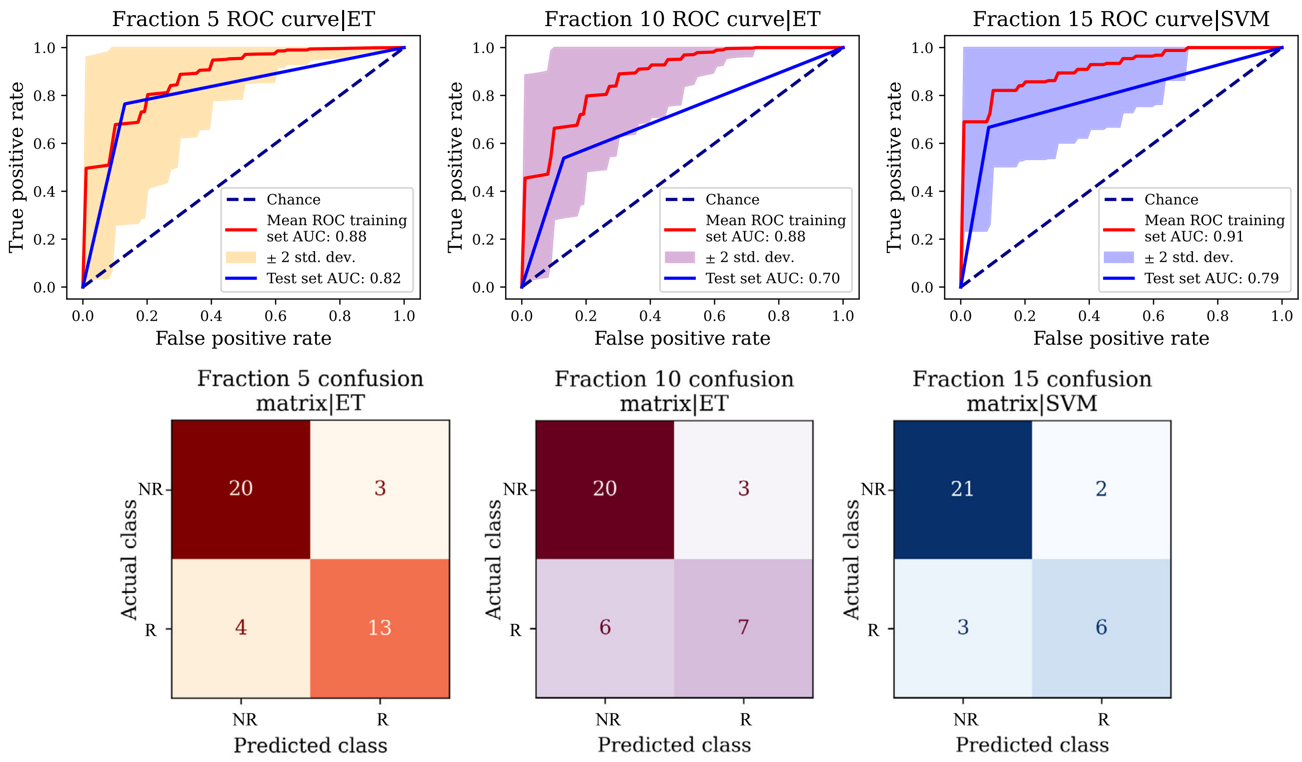

FIGURE 5 Fraction-specific model results. At the top are the ROC curves of the best model based on the training/validation CV results. At the bottom row, the confusion matrix results from the test set evaluation are shown. ET: ExtraTree Classifier and SVM: Support Vector Machine, with a sigmoid kernel. R stands for replanned and NR for non-replanned. CV, cross validation; ROC, receiver operating characteristic.

Furthermore, the definition of the metrics, along with the *HNGeoNatomyX* pipeline, will facilitate a potential analysis of Δ% values that may help create a standard guideline or criteria for replanning evaluation, such as the ones proposed by Weppler et al. (2018)[54] and Bak et al. (2022).[34] Thus, statistical approaches based on ROC curves or histograms of the metric variation may be a good approach for establishing clinical threshold values.[55]

Our results show the potential of early prediction based on the ability of the created metrics to describe the differences between the two classes. By comparing just the rate of change of the metrics, the null hypothesis test gave promising results starting with fractions 4 and 7, with p-value <0.05, which remained statistically significant for further fractions. These results corresponded to the first week of treatment.

Additionally, the incorporation of the metrics in an ML model analysis showed promising results in predicting whether replanning is needed, which could facilitate the RT treatment workflow for HNC patients. These models offer the potential to help clinical personnel determine as early as possible which patients will likely need treatment replanning, allowing timely intervention and improved workflow efficiency, and different models can be applied at each treatment fraction, incorporating the information available up to and including that fraction. The type of metrics selected by the automatic FS techniques (see Tables D, E, and F) included expected metrics with high predictive power of neck, mask, and PTV distance-related changes. Due to their high AUC in the univariate anaysis, we expected to find after the FS technique in our ML models the following metrics: $\bar{D}_{Body}$, CD_{Neck}, $\bar{B}_{mask}$, and A_{sub} (which were meant to quantify the mass loss) and $VI_{PTV} : LV_{Body}$ (related to PTV position with respect to the body). As shown in Tables D, E, and F, almost all of them were selected, with the exception of $\bar{D}_{Body}$; however, D_{Body} was selected, which also has a high AUC. Additionally, $\bar{B}_{mask}$, A_{sub}, and $VI_{PTV} : LV_{Body}$ were selected for the three specific ML fraction models, indicating that the created metrics appeared to be clinically relevant for the ML.

To the best of our knowledge, no other studies have done an ML analysis to predict RT replanning based on clinical data and geometrical metrics of the patient's shape over time. Given that our methodology includes information that radiation oncologists are familiar with in the clinic, the replanning prediction can be used for double verification. However, further analysis is still needed.

Since only geometrical metrics were selected by the FS techniques, and given that in our study sample, many patients' cancer subsites are located superior to the oropharynx (oral cavity, nasopharynx, nasal cavity, and sinus), incorporating additional metrics to describe only these subsites' changes could improve the quantification of mass loss. Furthermore, our study would benefit from evaluating metrics and ML models across different patient subgroups, such as HNC subsites, and by examining the differences between surgical cases (pre- and postoperative), as this distinction was not made in our current study due to information availability; however, it might also serve as a predictor.[56]

Regarding clinical implementation, once the models can be trained on a larger dataset and tested in an external cohort, this tool may potentially be integrated into treatment planning software to support decisions for RT treatment replanning. Including the automatic extraction pipeline in such software may help track important features and statistics over time, contributing to a standardized replan decision. Currently, this decision depends on individual radiation oncologists' and medical physicists' judgments, leading to larger variations in the decision. We hope that the development of ML models, such as the ones we developed for three specific time points, could predict and track changes over IMRT fractions. In the long term, we envision an automated workflow that would apply the appropriate model at the appropriate time.

Additional metrics, including OARs, can be further examined. This includes the movement of the parotid glands, as it has been shown that their spatial orientation appears to shift during treatment.[17] However, the contouring process of the parotid glands in each CBCT scan is currently time-consuming. To our knowledge, no automatic contouring tool in CBCTs exists yet. Furthermore, the inclusion of a metric to describe positional errors or mismatched spinal cord can also be analyzed.[57] Additionally, we acknowledge that the inclusion of dosimetric metrics would improve the clinical relevance of our ML model; however, since our goal was to ascertain if our geometrical metrics would potentially predict replanning, we did not include them in this study. However, doing so would be a valuable future extension of this work.

We acknowledge some limitations in our study. One of the main limitations is that some non-replanned patients considered in our study may have actually needed to be replanned, but they were not due to undocumented logistical reasons, such as consideration of the number of remaining fractions. Some of these cases might be reflected in the false positive cases shown in the confusion matrix. To overcome this type of limitation, a comprehensive analysis of the patients initially labeled as non-replanned needs to be reassessed. This could be done on a case-by-case basis and/or using clustering techniques.

Another limitation worth mentioning is that the metrics were extracted for all the fractions available simultaneously, which affects the body contour-cutting process to match the height of the bodies. Although this is correct for comparing the same region, to improve the input for ML models in a real-life scenario, this step should consider the contours up to the fraction to which the specific model corresponds. Another point to discuss is that, despite the efforts to avoid overfitting by performing a hold-out method, cross-validation, data standardization, and feature selection techniques, many created geometrical metrics may be colinear, which could lead to potential overfitting. This may be avoided in a future study by using Principal Component Analysis.[58]

Finally, it is important to mention that currently our models might not be generalizable since, even though our study sample covered a wide range of HNC subsites, it did not have significant samples for certain cases (salivary glands, nasal cavity, and hypopharynx), which might not be enough to provide confident results. However, with our developed and tested workflow and our code available under an open-source license, an analysis with a large, and ideally multicentre dataset, should be possible.

5 | CONCLUSIONS

We defined a number of geometrical metrics that describe the anatomical changes experienced by HNC patients during RT and demonstrated that they can be used to characterize and distinguish patients who will and will not require replanning, building an ML model. Our automatic metric calculation pipeline (*HNGeoNatomyX*) and associated ML models offer the potential to help streamline HNC patient resource management and clinical workflow in RT.

AUTHOR CONTRIBUTIONS

Odette Rios-Ibacache: Conceptualization; methodology and software; formal analysis; data curation; manuscript writing; manuscript editing. **James Manalad**: Conceptualization; methodology and software; manuscript editing. **Kayla O'Sullivan-Steben**: Formal analysis; data curation; manuscript editing. **Emily Poon**: Data curation; manuscript editing. **Luc Galarneau**: Formal analysis; manuscript editing. **Julia Khriguian**: Conceptualization; manuscript editing. **George Shenouda**: Conceptualization; manuscript editing. **John Kildea**: Conceptualization; formal analysis; manuscript writing; manuscript editing; supervision; funding acquisition. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS

This research was supported by a grant from the Rossy Cancer Network, the CREATE grant in Responsible Health and Healthcare Data Science (SDRDS) from the Natural Sciences and Engineering Research Council, the Fonds de recherche du Quebec–Santé dual chair in AI and Digital Health of John Kildea, and a studentship award to Odette Rios-Ibacache from the Research Institute of the MUHC (RI-MUHC). The authors gratefully acknowledge Ackeem Joseph; Aixa X. Andrade Hernandez; and Victor Matassa for their valuable contributions in sourcing the data used in this study.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The automatic extraction pipelines for calculating the metrics and the semi-automatic treatment mask contouring algorithm that support the findings of this study are available under an open source license on kildealab GitHub in the HNGeoNatomyX section: github.com/kildealab/HNGeoNatomyX.

ETHICAL APPROVAL STATEMENT

This retrospective study was approved by the McGill University Health Centre's Research Ethics Board (REB). All work was conducted in accordance with REB guidelines.

ORCID

Odette Rios-Ibacache
https://orcid.org/0009-0003-9254-5592
Kayla O'Sullivan-Steben
https://orcid.org/0009-0002-5993-7968
Luc Galarneau
https://orcid.org/0000-0003-1871-9331
John Kildea https://orcid.org/0000-0002-7084-1425

REFERENCES

1. Amaral MN, Faísca P, Ferreira HA, Gaspar MM, Reis CP. Current insights and progress in the clinical management of head and neck cancer. *Cancers (Basel)*. 2022;14(24):6079.
2. Grégoire V, Langendijk JA, Nuyts S. Advances in radiotherapy for head and neck cancer. *J Clin Oncol*. 2015;33(29):3277-3284. PMID: 26351354 http://doi.org/10.1200/JCO.2015.61.2994
3. Tsai PC, Liu YC, Li TS, et al. Clinical effect of moisturized skin care on radiation dermatitis of head and neck cancer. *In Vivo*. 2023;37(6):2776–2785.
4. Larsson M, Hedelin B, Athlin E. Lived experiences of eating problems for patients with head and neck cancer during radiotherapy. *J Clin Nurs*. 2003;12(4):562–570.
5. Sroussi HY, Epstein JB, Bensadoun RJ, et al. Common oral complications of head and neck cancer radiation therapy: mucositis, infections, saliva change, fibrosis, sensory dysfunctions, dental caries, periodontal disease, and osteoradionecrosis. *Cancer Med*. 2017;6(12):2918–2931.
6. Barker JL, Garden AS, Ang KK, 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(4):960–970.
7. Jham BC, Silva Freire dAR. Oral complications of radiotherapy in the head and neck. *Braz J Otorhinolaryngol*. 2006;72(5):704–708.
8. Robar JL, Day A, Clancey J, et al. Spatial and dosimetric variability of organs at risk in head-and-neck intensity-modulated radiotherapy. *Int J Radiat Oncol Biol Phys*. 2007;68(4):1121–1130.
9. Figen M, Çolpan Öksüz D, Duman E, et al. Radiotherapy for head and neck cancer: Evaluation of triggered adaptive replanning in routine practice. *Front Oncol*. 2020;10:579917.
10. Chinnery TA, Lang P, Nichols AC, Mattonen SA. Predicting the need for a replan in oropharyngeal cancer: A radiomic, clinical, and dosimetric model. *Med Phys*. 2024;51(5):3510–3520.
11. Brown E, Owen R, Harden F, et al. Predicting the need for adaptive radiotherapy in head and neck cancer. *Radiother Oncol*. 2015;116(1):57–63.
12. Beek vS, Jonker M, Hamming-Vrieze O, 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. Zhou X, Wang W, Zhou C, et al. Long-term outcomes of replanning during intensity-modulated radiation therapy in patients with nasopharyngeal carcinoma: An updated and expanded retrospective analysis. *Radiother Oncol*. 2022;170:136-142. https://doi.org/10.1016/j.radonc.2022.03.007
14. Bhide SA, Davies M, Burke K, 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(5):1360–1368.
15. Fiorentino A, Caivano R, Metallo V, et al. Parotid gland volumetric changes during intensity-modulated radiotherapy in head and neck cancer. *Br J Radiol*. 2012;85(1018):1415–1419.
16. Wang W, Yang H, Hu W, et al. Clinical study of the necessity of replanning before the 25th fraction during the course of intensity-modulated radiotherapy for patients with nasopharyngeal carcinoma. *Int J Radiat Oncol Biol Phys*. 2010;77(2):617–621.
17. Morgan HE, Sher DJ. Adaptive radiotherapy for head and neck cancer. *Cancers Head Neck*. 2020;5(1):1.
18. Walls R, McLaughlin C. Comparison and breakdown of cost in head and neck cancer radiation therapy and adaptive replanning. *Int J Radiat Oncol*Biol*Phys*. 2023;117(2, Supplement):e632. ASTRO 2023: 65th Annual Meeting. https://doi.org/10.1016/j.ijrobp.2023.06.2029
19. Heukelom J, Fuller CD. Head and neck cancer adaptive radiation therapy (ART): Conceptual considerations for the informed clinician. *Semin Radiat Oncol*. 2019;29(3):258–273.
20. Volpe S, Pepa M, Zaffaroni M, et al. Machine learning for head and neck cancer: A safe bet?-A clinically oriented systematic review for the radiation oncologist. *Front Oncol*. 2021;11:772663.
21. Brown E, Owen R, Harden F, et al. Head and neck adaptive radiotherapy: Predicting the time to replan. *Asia Pac J Clin Oncol*. 2016;12(4):460–467.
22. Delaby N, Barateau A, Chiavassa S, et al. Practical and technical key challenges in head and neck adaptive radiotherapy: The GORTEC point of view. *Phys Med*. 2023;109(102568):102568.
23. Oommen T, Baise LG, Vogel RM. Sampling bias and class imbalance in maximum-likelihood logistic regression. *Math Geosci*. 2011;43(1):99–120.
24. Mata dS, Ferreira J, Nicolás I, et al. P16 and HPV genotype significance in HPV-associated cervical cancer-A large cohort of two tertiary referral centers. *Int J Mol Sci*. 2021;22(5):2294.
25. Vangelov B, Kotevski DP, Williams JR, Smee RI. The impact of HPV status on weight loss and feeding tube use in oropharyngeal carcinoma. *Oral Oncol*. 2018;79:33–39.
26. Zhu H, Amdur R, Lu B, Mendenhall W. Adaptive replanning for HPV-associated N2b oropharyngeal squamous cell carcinoma in response to anatomic changes during radiation therapy. *Int J Radiat Oncol*Biol*Phys*. 2018;100:1358-1359. https://doi.org/10.1016/j.ijrobp.2017.12.135
27. Park D, Jain S, Quay-De La Vallee Z, Huber K, O'Leary M, Farag AM. Influence of smoking history on the perception of radiation-induced dysgeusia/hypogeusia in patients with head and neck cancer. *Eur Arch Otorhinolaryngol*. 2021;278(8):2993–3001.
28. Zhao L, Wan Q, Zhou Y, Deng X, Xie C, Wu S. The role of replanning in fractionated intensity modulated radiotherapy for nasopharyngeal carcinoma. *Radiother Oncol*. 2011;98(1):23–27.

29. Mason D, scaramallion, mrbean-bremen, et al. pydicom/pydicom: pydicom v2.4.4. 2023.
30. Brivio D, Hu YD, Margalit DN, Zygmanski P. Selection of head and neck cancer patients for adaptive replanning of radiation treatment using kV-CBCT. *Biomed Phys Eng Express*. 2018;4(5):055009. https://doi.org/10.1088/2057-1976/aad546
31. Williams F. Point Cloud Utils. 2022. https://www.github.com/fwilliams/point-cloud-utils
32. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17:261–272. https://doi.org/10.1038/s41592-019-0686-2
33. Sullivan CB, Kaszynski A. PyVista: 3D plotting and mesh analysis through a streamlined interface for the visualization toolkit (VTK). *J Open Source Softw*. 2019;4(37):1450. https://doi.org/10.21105/joss.01450
34. Bak B, Skrobala A, Adamska A, 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(5):722.
35. Kumar A, Soni TP, Patni N, et al. An observational dosimetric study in definitively treated primary head and neck cancers: To assess the effect of weight loss and change in lateral neck dimensions on the difference between dose planned and received by the parotid(s) and correlation with adaptive radiation therapy. *Adv Radiat Oncol*. 2024;9(5):101446.
36. Ahn PH, Chen CC, Ahn AI, et al. Adaptive planning in intensity-modulated radiation therapy for head and neck cancers: single-institution experience and clinical implications. *Int J Radiat Oncol Biol Phys*. 2011;80(3):677–685.
37. El-Sayed S, Yemchuk S, Broomfield J, Bahn J. Is percentage of weight loss predictive of the need for re-planning of patients with head and neck cancer treated with IMRT radiotherapy? Results of a prospective study. *Int J Radiat Oncol Biol Phys*. 2011;81(2):S540-S541. https://doi.org/10.1016/j.ijrobp.2011.06.843
38. Stauch Z, Zoller W, Tedrick K, et al. An evaluation of adaptive planning by assessing the dosimetric impact of weight loss throughout the course of radiotherapy in bilateral treatment of head and neck cancer patients. *Med Dosim*. 2020;45(1):52–59.
39. Andrade-Hernandez A, Kildea J. *Exploration of Potential Parameters that Influence when to Replan Head and Neck Cancer Patients during Radiotherapy*, 2020.
40. Armstrong RA. When to use the Bonferroni correction. *Ophthalmic Physiol Opt*. 2014;34(5):502–508.
41. Mason S, Graham N. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical signié cance and interpretation. *Q J R Meteorol Soc*. 2002;128:2145-2166. https://doi.org/10.1256/003590002320603584
42. Xie D, Xu F, Zhu W, et al. Delta radiomics model for the prediction of progression-free survival time in advanced non-small-cell lung cancer patients after immunotherapy. *Front Oncol*. 2022;12:990608.
43. Cousin F, Louis T, Dheur S, et al. Radiomics and delta-radiomics signatures to predict response and survival in patients with non-small-cell lung cancer treated with immune checkpoint inhibitors. *Cancers (Basel)*. 2023;15(7):1968.
44. Avgousti R, Antypas C, Armpilia C, et al. Adaptive radiation therapy: When, how and what are the benefits that literature provides?. *Cancer Radiother*. 2022;26(4):622–636.
45. Gan Y, Langendijk JA, Oldehinkel E, Lin Z, Both S, Brouwer CL. Optimal timing of re-planning for head and neck adaptive radiotherapy. *Radiother Oncol*. 2024;194(110145):110145.
46. Kuhn M, Johnson K. *Applied predictive modeling*. New York, NY: Springer; 2019.
47. Chen Q, Meng Z, Liu X, Jin Q, Su R. Decision variants for the automatic determination of optimal feature subset in RF-RFE. *Genes (Basel)*. 2018;9(6):301.
48. Darst BF, Malecki KC, Engelman CD. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. *BMC Genet*. 2018;19(Suppl 1):65.
49. Fan X, Xie N, Chen J, et al. Multiparametric MRI and machine learning based radiomic models for preoperative prediction of multiple biological characteristics in prostate cancer. *Front Oncol*. 2022;12:839621.
50. Hadjiiski L, Cha K, Chan HP, et al. AAPM task group report 273: Recommendations on best practices for AI and machine learning for computer-aided diagnosis in medical imaging. *Med Phys*. 2023;50(2):e1–e24.
51. Nuyts S, Bollen H, Eisbruch A, et al. Adaptive radiotherapy for head and neck cancer: Pitfalls and possibilities from the radiation oncologist's point of view. *Cancer Med*. 2024;13(8):e7192.
52. Hu YC, Tsai KW, Lee CC, et al. Which nasopharyngeal cancer patients need adaptive radiotherapy?. *BMC Cancer*. 2018;18(1):1234.
53. Chen AM, Daly ME, Cui J, Mathai M, Benedict S, Purdy JA. Clinical outcomes among patients with head and neck cancer treated by intensity-modulated radiotherapy with and without adaptive replanning. *Head Neck*. 2014;36(11):1541–1546.
54. Weppler S, Quon H, Banerjee R, Schinkel C, Smith W. Framework for the quantitative assessment of adaptive radiation therapy protocols. *J Appl Clin Med Phys*. 2018;19(6):26–34.
55. Hassanzad M, Hajian-Tilaki K. Methods of determining optimal cut-point of diagnostic biomarkers with application of clinical data in ROC analysis: an update review. *BMC Med Res Methodol*. 2024;24(1):84.
56. Gourin CG, Couch ME, Johnson JT. Effect of weight loss on short-term outcomes and costs of care after head and neck cancer surgery. *Ann Otol Rhinol Laryngol*. 2014;123(2):101–110.
57. Wang X, Lu J, Xiong X, et al. Anatomic and dosimetric changes during the treatment course of intensity-modulated radiotherapy for locally advanced nasopharyngeal carcinoma. *Med Dosim*. 2010;35(2):151–157.
58. Vasicek D. Artificial intelligence and machine learning: Practical aspects of overfitting and regularization. *Inf Serv Use*. 2020;39(4):281–289.

How to cite this article: Rios-Ibacache O, Manalad J, O'Sullivan-Steben K, et al. Quantification of head and neck cancer patients' anatomical changes during radiotherapy: Toward the prediction of replanning need. *J Appl Clin Med Phys*. 2026;27:e70465. https://doi.org/10.1002/acm2.70465