Clinical and Translational Radiation Oncology 39 (2023) 100590

Contents lists available at ScienceDirect

Clinical and Translational Radiation Oncology

journal homepage: www.sciencedirect.com/journal/clinical-and-translational-radiation-oncology

Artificial intelligence to predict outcomes of head and neck radiotherapy

Chulmin Bang [a,*], Galaad Bernard [a], William T. Le [b,c], Arthur Lalonde [a,b,d], Samuel Kadoury [b,c], Houda Bahig [a,b]

[a] Centre Hospitalier de l'Université de Montréal, Montreal, QC, Canada
[b] Centre de recherche du Centre Hospitalier de l'Université de Montréal, Montreal, QC, Canada
[c] Polytechnique Montréal, Montreal, QC, Canada
[d] Université de Montréal, Montreal, QC, Canada

ARTICLE INFO

Keywords:
Head and neck cancer
Artificial intelligence
Machine learning
Radiomic
Predictive modeling
Cancer outcomes

ABSTRACT

Head and neck radiotherapy induces important toxicity, and its efficacy and tolerance vary widely across patients. Advancements in radiotherapy delivery techniques, along with the increased quality and frequency of image guidance, offer a unique opportunity to individualize radiotherapy based on imaging biomarkers, with the aim of improving radiation efficacy while reducing its toxicity. Various artificial intelligence models integrating clinical data and radiomics have shown encouraging results for toxicity and cancer control outcomes prediction in head and neck cancer radiotherapy. Clinical implementation of these models could lead to individualized risk-based therapeutic decision making, but the reliability of the current studies is limited. Understanding, validating and expanding these models to larger multi-institutional data sets and testing them in the context of clinical trials is needed to ensure safe clinical implementation. This review summarizes the current state of the art of machine learning models for prediction of head and neck cancer radiotherapy outcomes.

Introduction

Global approximate incidence of head and neck cancer (HNC) is 880 000 patients each year worldwide [1]. As a core therapeutic option, radiotherapy (RT) is being used in 75 % of cases, combined with other treatment modalities such as chemotherapy or surgery [2]. Current HNC RT is associated with high rates of toxicity as well as adverse impacts on patients' quality of life [3]. HNC encompasses a heterogeneous group of cancers originating from various subsites that are associated with various risk factors including viral infections [4], tobacco and alcohol use [5]. In addition to known heterogeneity, RT efficacy and tolerance also vary across apparently similar patients presenting with the same cancer subtype, anatomical stage and apparent risk factors [6]. Patient-specific clinical, radiological and biological factors are thought to drive these individual cancer outcomes. To date, HPV-positive status is known to be a favorable predictor of response to treatment in oropharyngeal

Abbreviations: ADASYN, adaptive synthetic sampling; AI, artificial intelligence; ANN, artificial neural network; AUC, Area Under the ROC Curve; BMI, body mass index; CART, Classification and Regression Tree; CBCT, cone-beam computed tomography; C-Index, concordance index; CIFE, conditional informax feature extraction; CNN, convolutional neural network; CRT, chemoradiation; CT, computed tomography; DL, deep learning; DT, Decision Tree; DM, distant metastasis; DSC, Dice Similarity Coefficient; DSS, clinical decision support systems; DVH, Dose-volume histogram; GANs, Generative Adversarial Networks; GB, Gradient boosting; GPU, graphical process units; HNC, head and neck cancer; HPV, human papillomavirus; HR, hazard ratio; IAMB, incremental association Markov blanket; IBDM, image based data mining; IBMs, image biomarkers; IMRT, intensity-modulated RT; KNN, k nearest neighbor; LLR, Local linear forest; LR, logistic regression; LRR, loco-regional recurrence; ML, machine learning; MIFS, mutual information based feature selection; NPC, nasopharynx; MRMR, Minimum redundancy feature selection; MRI, Magnetic resonance imaging; *N*-MLTR, Neural Multi-Task Logistic Regression; NTCP, Normal Tissue Complication Probability; OPC, oropharyngeal cancer; ORN, osteoradionecrosis; OS, overall survival; PCA, Principal component analysis; PET, Positron emission tomography; PG, parotid glands; PLR, Positive likelihood ratio; PM, pharyngeal mucosa; PTV, Planning target volumes; PreSANet, deep preprocessor module and self-attention; QUANTEC, Quantitative Analyses of Normal Tissue Effects in the Clinic; RF, random forest; RFC, random forest classifier; RFS, recurrence free survival; RLR, Rigid logistic regression; RSF, random survival forest; RT, radiotherapy; RTLI, radiation-induced temporal lobe injury; RRF, Regularized random forest; SDM, shared decision making; SMG, submandibular glands; SMOTE, synthetic minority over-sampling technique; STIC, sticky saliva; SVC, support vector classifier; SVM, support vector machine; XGBoost, extreme gradient boosting.

* Corresponding author at: Centre Hospitalier de l'Université de Montréal, 3840 Rue Saint-Urbain, Montréal, QC H2W 1T8, Canada.
E-mail address: chul.min.bang@umontreal.ca (C. Bang).

https://doi.org/10.1016/j.ctro.2023.100590
Received 13 January 2023; Received in revised form 28 January 2023; Accepted 28 January 2023
Available online 31 January 2023

cancer and have led to the concept of de-escalation treatments, but which can not be used outside of clinical trials as studies have been inconclusive [18]. Currently, there are no other biomarkers used to guide decisions in HNC RT.

Over the last decades, the refinement of RT techniques, along with the parallel developments in image guidance for better delineation and tumor localization over the course of treatment [7–8], have been associated with an increasing interest for individualized RT with the aim of increasing or maintaining tumor control and reducing radiation toxicity [7]. Machine learning (ML) consists in the analysis of large amounts of empirical data using computational algorithms and leading to automatic extraction of discriminative features and learning of complex patterns [9–10]. Developments in artificial intelligence (AI), particularly ML and deep learning (DL), have led to significant enthusiasm for the concept of "rapid learning health system", whereby decision-making would be individualized based on analysis of large patient cohorts [11]. AI offers a unique opportunity for the development of predictive models that can help stratify individual patient's risk and guide therapeutic decisions for optimal patients' outcomes and quality of life in HNC RT. Herein, we review the current stance of the role of AI in predicting toxicity and therapeutic outcomes in HNC patients treated with RT.

Radiation oncology and artificial intelligence

The term "Artificial intelligence" first appeared at the Dartmouth Scientific Conference in 1956 [12]. ML is a subfield of AI that uses an algorithm to find patterns within data. DL is a subcategory of ML based on the use of neural networks and representation learning (Fig. 1).

Since the success of computer vision based analysis due to the increase in access to graphical process units (GPU), the field of medical image analysis has seen rapid benefits from ML approaches, which contributed in advancing the fast growing field of adaptive RT. With the increasing digitization and standardization of image acquisition and storage processes, the primary drawback of ML methods - the large training data requirements - continues to diminish. More specifically, recent studies focused on opportunities to use *big data* as decision support by predicting tumor response and toxicity outcomes [19–20]. ML-based risk stratification could support decisions such as RT or systemic therapy intensification (or de-intensification), or even guide prophylactic measures in prevention of expected toxicity.

Radiomics (Fig. 2) is a research field which consists in the extraction of a large number of quantitative, hand-crafted features from different medical images (CT; MRI; PET-CT) - that cannot be manually deciphered by clinicians [21], and which will be then integrated to data-characterization algorithms [22]. Although pathology can be gold standard, it is not feasible to biopsy every node or every area of the tumor and radiomics offers the promise of characterizing different areas of the cancer [23]. Radiomics offer a particular potential in HNC RT as an immense volume of different imaging modalities are gathered routinely. Most HNC patients begin with initial diagnostic and staging images including diagnostic and/or planning computed tomography (CT) and magnetic resonance imaging (MRI). Positron emission tomography (PET) is also frequently part of routine diagnostic (or planning) imaging in HNC [24–25]. During RT, daily CBCT (cone-beam computed tomography) serves for patient positioning further contributing to the quantity of available anatomic imaging [26]. More recently, the MR-Linac, which is available in an increasing number of institutions, has allowed for daily online MRI, therefore further improving image guidance [27]. After RT completion, follow-up typically includes CT, MRI or PET, based on different factors and guidelines [28].

Classification models are used to cluster data into different groups by approximating a mapping from a different set of inputs to discrete a set of outputs, has been mainly used as a predictive ML algorithm for HNC. The main algorithms are logistic regression (LR), decision tree, random forest (RF), support vector machine (SVM), k nearest neighbor (KNN), naives Bayes and artificial neural network. These traditional models have allowed better understanding of medical images, but with the major disadvantage of the time consuming step involving extracting and selecting important image features [35]. To address this issue, *convolution neural network* (CNN), a DL model that automatically learns, extracts and selects important features have shown promise in imaging classification tasks [36]. For any AI models, including classifiers, *overfitting* is a common problem when an algorithm is too focused on the training data set that it can not be successfully applied to a new data set. In order to verify if the model is not overfitting, a k-fold cross validation could be done by randomly partitioning a data set into k mutually exclusive subsets. Different metrics can be used in order to evaluate a model's performance. These include sensitivity, specificity, accuracy, precision and recall. Another measurement could be done with the Area Under the ROC Curve (AUC) ranging from 0 to 1 (0 for a model's predictive performance is 100 % wrong and 1, as its performance is 100 % correct). The calculation of the concordance index (C-index) could also be measured for predictive performance, on a scale between zero and one with 0.5 indicating no prognostic value.

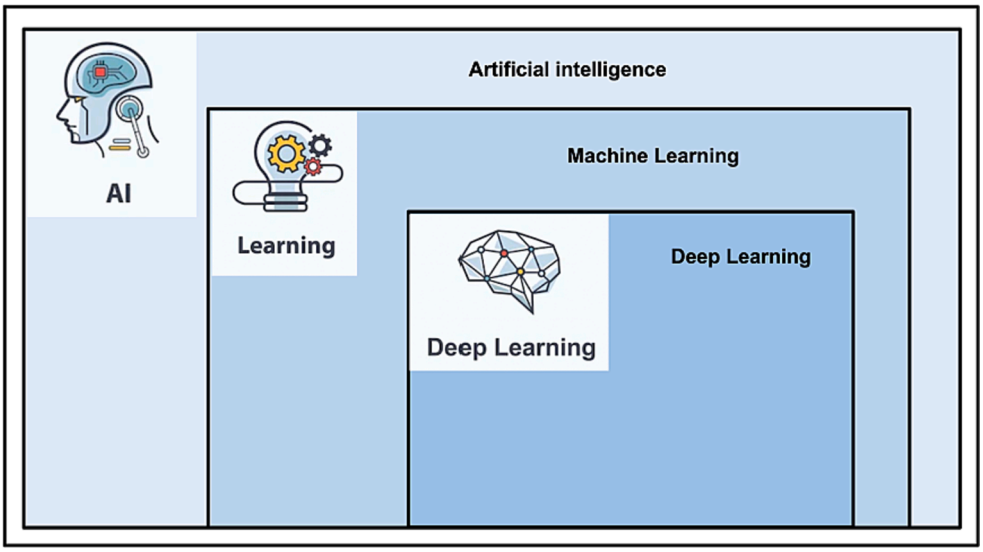

Fig. 1. Various levels of Artificial intelligence, including machine learning and deep learning.

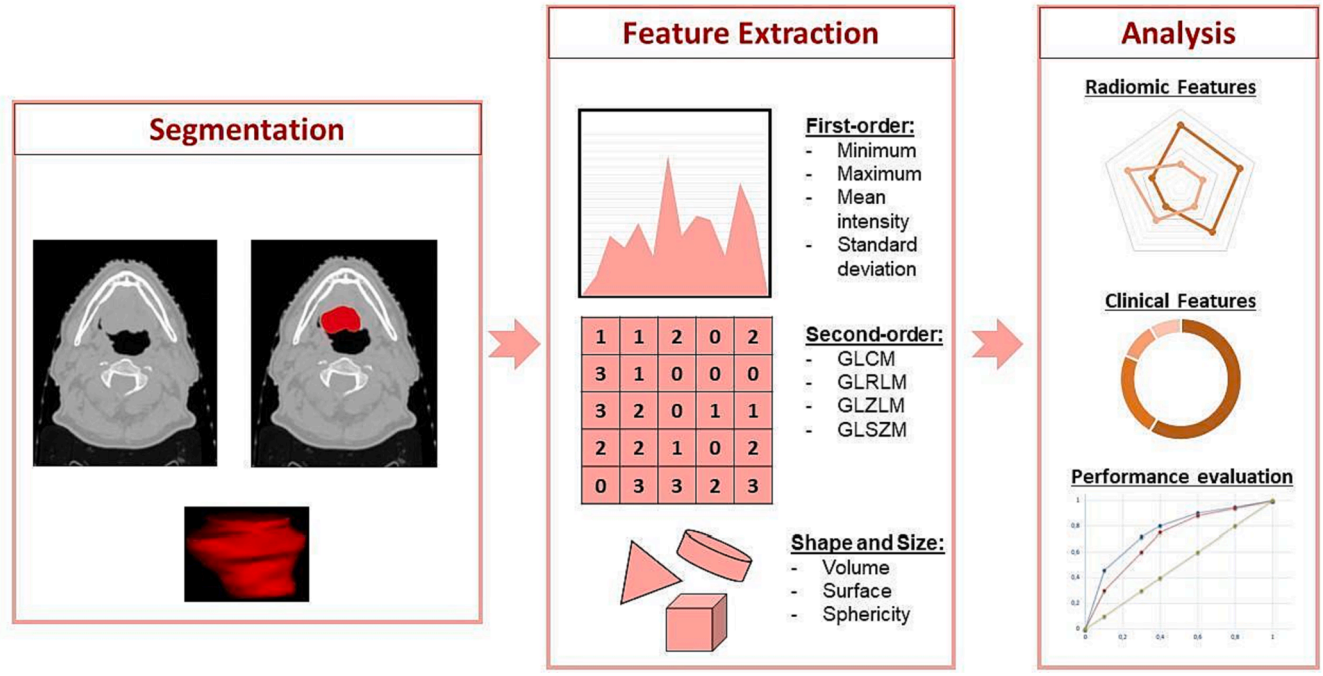

Fig. 2. Schematic illustration of radiomic pipeline.

Table 1
Summary list of potential use of AI algorithms in different steps in HNC RT.

Steps	Clinical Use	Algorithms
1) Diagnosis/ Initial evaluation	Clinical/Pathological/ radiological data processing	SVM, LR, RF, DT, KNN, Bayesian, Linear Discriminant Analysis, DL and combination of ML and DL
2) Treatment Decision Making	Treatment Decision Aid	DSS, SDM
	Outcome Prediction	SVM, LR, Bayesian, neural network, decision trees and combination
	Toxicities Prediction	Bayes, LR, KNN, SVM, SVC, RF, XGBoost PLR, RFC, IBDM, CART, CNN
3) Simulation	Synthetic Imaging Generation	Fuzzy c-means clustering, CNNs and GANs
4) Treatment Planning	Auto-segmentation	DSC, CNN DD-Resnet
	Dosimetric Optimization	Unet, DenseNet, GANs
5) Treatment delivery	Image Guidance and Motion Management	CNN, Bayesian, SVM,
	CBCT quality improvement	CNN

Abbreviations: CBCT, cone-beam computed tomography; CNN, convolutional neural network; CART, Classification and Regression Tree; DL, deep learning; DT, Decision Tree; DSC, Dice Similarity Coefficient; DSS, clinical decision support systems; KNN, k nearest neighbor; LR, logistic regression;ML, machine learning; GANs, Generative Adversarial Networks; IBDM, image based data mining; RF, random forest; RFC,random forest classifier; SDM, shared decision making; SVM, support vector machine; SVC, support vector classifier.

Prediction of toxicity

Different Normal Tissue Complication Probability (NTCP) models have been proposed in order to help understand and estimate the risk of different toxicities. The common traditional models are the Lyman-Kutcher-Burnman models based on simplified characterization of the radiation-dose response for each anatomy [37]. For different toxicities, some studies have shown good validation [38]. In clinical practice, radiation oncologists routinely use the Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC) dose–response guidelines to determine tolerance dose to organs at risk [39]. However, the predictive power of currently models is largely limited by the lack of integration of patient-specific risk factors [40 41] and omission of the inclusion of radiomic features despite studies having shown their relevance [38]. In order to account for these issues and update the current NTCP models, different ML methods have been studied [43]. Table 2 summarizes the current literature on AI predictive models for toxicities.

As xerostomia is known to be a common major early- and long-term toxicity of HNC RT, different studies reported on radiomic features from salivary glands, namely the parotid glands (PGs) or submandibular glands (SMGs). Gabry et al [44] compared the predictive performance of seven classification algorithms, six feature selection methods, and ten data cleaning/class balancing techniques by integrating retrospective dosiomic, radiomic and demographic data from 153 HNC patients treated with definitive RT. Important features such as dosimetric shape and gradients, PG volume and eccentricity, and DVH spread were identified. Their MLs based models performed better compared to the traditional NTCP models based on mean dose to PG only. Jiang et al [45] tested three different algorithms applied to a data set including clinical and radiomic features such as voxels dose in the PGs and SMGs from 427 HNC patients treated with definitive RT. Similar to. Gabry et al, the authors were able to observe that specific dose patterns across the subvolumes in both organs were an important predictive feature and their ridge logistic regression model had the best performance with AUC of 0.70. Dijk et al [46] used the CT based image biomarkers (IBMs) of the PGs and SMGs from 249 HNC patients treated with definitive RT in order to improve the NTCP predictive models for sticky saliva (STIC) and moderate-severe xerostomia 12 months. For both toxicities, pre-selection through a lasso regularization identified different important radiomic features and the predictive performance of NTCP models were stronger when IBMs were added (AUC 0.74 vs 0.77). Beasley et al [47] also used ML applied to the image based data mining (IBDM) in order to identify clusters of dose distribution voxels involved in radiation induced trismus. From clinical and radiomic data set from 86 HNC patients focusing on the dose distributions within the anatomy of interest, different features investigated in a multivariable analysis and an internal/external validation demonstrated the importance of dose patterns within anatomy to predict trismus.

Table 2
AI predictive models for toxicities.

Authors	Endpoints	Number of HNC cohort	Algorithms	Performance	Important features
Dean et al. (2016) [53]	Mucositis	351	PLR, SVC, RFC	RFC with AUC 0.71	Volumes of oral cavity receiving intermed - high dose
Dijk et al. (2016) [46]	Xerostomia and sticky saliva at 12 months	249	LASSO regularisation	AUC 0.77	IBMs
Pota et al. (2017) [95]	PG shrinkage and Xerostomia at 12 months	37	fuzzy classification	AUC 0.86 and 0.79	Final volume PG shrinkage
Dean et al. (2017) [49]	Dysphagia	173	PLR, SVC, RFC	RFC with AUC 0.71	PM receiving >1 Gy/fraction
Cheng et al. (2017) [52]	WL during RT and EOT	391	CART	AUC 0.773 and 0.821	Dose in OARs, oral intake, N stage, pain, nausea
Beasley et al. (2018) [47]	Trismus	86	Linear regression	Rs of −0.45	IBDM clusters in ipsilateral masseter
Gabry s et al. (2018) [44]	Xerostomia at different timelines	153	7 classifiers and 6 feature selectors	AUC 0.74–0.88	Dose gradient in PGs, PG volume, PG eccentricity
Jiang et al. (2018) [45]	Xerostomia 3 months post RT	427	RLR, LLR, RF	AUC 0.70	Dose pattern in PG/SMG
Reddy et al. (2019) [50]	Hospitalization Feeding tube WL	2121	RF, GB, LR	AUC 0.640-0.751	–
Wojcieszynski et al (2019) [51]	Grade > 3 toxicity (90and180days)	437	PLR, RF, XGBoost	C-statistic 0.65 and 0.63	PTV integral dose and integral dose out of PTV
Zhang et al. (2020) [54]	RTLI post RT in different timelines	242 NPC	RF	AUCs 0.830, 0.773 and 0.716	Features from medial temporal lobe
Humbert-Vidan et al. (2022) [55]	ORN	140	3D CNN, DenseNet 21	AUC 0.71	Clinical dosimetric distribution

Abbreviations: AUC, Area Under the ROC Curve; CART, Classification and Regression Tree; GB, Gradient boosting; IBMs, image biomarkers; LLR, Local linear forest; LR, logistic regression; PLR, Positive likelihood ratio; PG, parotid glands; PTV, Planning target volumes; RLR, Rigid logistic regression; RF, random forest; RFC, random forest classifier; SMG, Submandibular glands; SVC, support vector classifier; XGBoost, extreme gradient boosting; 3D, three dimensions.

Dysphagia is also a common toxicity among patients treated with RT for HNC causing a major impact in patients' quality of life [48]. In order to improve the current predictive NTCP model for dysphagia, Dean et al [49] incorporated spatial dose metrics in different ML models with a prospective data set of 173 HNC patients including clinical and dosimetric features focusing on pharyngeal mucosa (PM). Their RFC model with a highest AUC of 0.71 identified that the volume, length and circumference of PM receiving 1 Gy/fraction and higher were strongly associated with the risk of dysphagia. More recently, Reddy et al [50] used a different data set from 2,121 HNC patients in order to compare predictive performance of three different classifiers for unplanned hospitalizations, feeding tube placement and significant weight loss. This method identified over 700 treatment-related and clinical variables, and achieved AUC values of up to 0.64, 0.75, and 0.75 for RF, gradient boosting, and LR, respectively. Wojcieszynski et al [51] compared the predictive performance of three ML methods on a prospective data set of 437 HNC patients treated with definitive chemoradiation (CRT). Their RF model yielded moderate success for toxicity at 90 and 180 days with c-static of 0.65 and 0.63, respectively. From this study, higher integral doses outside of the target volume, target volume integral dose, body mass index (BMI) and age were important factors associated with increased grade 3 + toxicity. Cheng et al [52] used a Classification and Regression Tree model in HNC patients treated with definitive RT by using demographic, dosimetric and clinical data from 391 patients in order to predict weight loss $\geq$5 kg at 3 months post-RT. Two models were built, one during the RT planning and one at the end of the treatments. When additional treatment-related data was added to each model, the predictive performance was improved, with an AUC of 0.77 and 0.82, respectively.

Several other radiation-induced toxicities have been the focus of individual studies. Mucositis was the focus of a study by Dean et al [53], where clinical, dose-volume and spatial dose metrics data from 317 HNC patients were used to build predictive ML models for severe acute mucositis. Among different models tested, the discriminative performance was not improved with the additional spatial dose metrics. Important features were in the range *V80-V220* and the most important feature was the *V220*. Interestingly, in contrast to the RTOG guidelines focusing on mean dose to the oral cavity, the authors therefore identified that the strongest feature associated with severe acute mucositis was the volume receiving intermediate and high doses. To predict radiation-induced temporal lobe injury, Zhang et al [54] used retrospective clinical and CT and MRI based radiomic data from 242 nasopharynx (NPC) patients treated with definitive RT. Different radiomic features were first extracted from the medial temporal lobe regions. RF predictive model showed strong predictive performance in three subsequent radiological follow-ups preceding the onset of radiation-induced temporal lobe injury with the mean AUCs of between 0.71 and 0.83. More recently, Humbert-Vidan et al [55] used retrospective demographic, clinical and dosimetric data (3D dose distribution map) of 140 HNC patients (70 patients with ORN and 70 patients as control group) in order to compare the predictive performance of a 3D densely-connected 121-layer convolutional neural network (CNN) model with a DVH based RF model. The 3D DenseNet121 CNN model had better performance with an average AUC of 0.71 (0.64–0.79), compared to 0.65 (0.57–0.73) for the RF model.

There was a general trend of increasing feature importance with increasing dose and feature importance was also high for RT dose metrics in the range V80 – V220.

Prediction of cancer control outcomes

Current predictive models for HNC are mainly based on the TNM (Primary tumor, regional lymph nodes and distant metastasis) staging [56], which guides oncologists in the selection of the appropriate therapeutic options for patients. In oropharyngeal cancer (OPC), other prognostic variables established in the context of clinical trials, namely human papillomavirus (HPV) status and tobacco smoking pack-years, have largely been integrated in the clinical practice. While patients with HPV-associated OPC generally have an improved prognosis, we now recognize that a subset of these patients present a highly aggressive behavior [57]. There is therefore a general concern that safe treatment de-escalation should not jeopardize the chance of cure of these patients, and that more reliable tools to predict tumor behavior are greatly needed. In recent years, several predictive AI models to better predict cancer control outcomes in HNC have been published and are presented in Table 3.

To better stratify survival outcomes, Tseng et al [58] evaluated the elastic net penalized Cox proportional hazards regression-based risk stratification model in operated oral cavity cancers. The authors integrated clinicopathologic and genomic data from 334 patients with

Table 3
Different AI predictive models for outcomes.

Authors	Patient population	Predicted outcomes	Number of patients	Algorithms	Performance	Important features
Parmar et al. (2015) [63]	HNC SCC	3Y OS	206	13 feature selectors and 11 ML classifiers	3 feature selectors (MRMR, MIFS and CIFE) with AUC 0.66–0.69 3 classifier (RF, NN and BY) with AUC 0.61–0.67	CT based radiomic features
Jiang et al. (2015) [62]	NSC	OS	347	SVM	HR 3.45	Combination of CRT
Li et al. (2018) [65]	NSC	LR	306	PCA, ANN, KNN SVM	ANN with accuracy of 0.812	MRI based radiomic features
Zdilar et al. (2018) [64]	OPC	OS RFC	529	MRMR, Wilcoxon, RF, RReliefF, RRF, IAMB, RSF, PCA	RF selectors AUC 0.75 and C-index 0.68	CT based radiomic features
Fujima et al. (2019) [96]	Sinonasal SCC	LF	36	SVM	Accuracy of 0.96	MRI based radiomic features
Wu et al. (2019) [97]	OPC	DMFS	140	RSF	C-index 0.73	Max distance between nodes and tumor-nodes
Zhou et al. (2019) [67]	NPC	DM	176	PyRadiomics features extraction, features selections (Mann-Whitney *U* test, mRMR, Lasso), LR	AUC 0.827 (training group) and 0.792 (validation group)	MRI based radiomic features
Tseng et al. (2020) [58]	Oral cavity	Survival (Cancer specific and loco-regional recurrence free)	334	Elastic Net Penalized Cox Proportional Hazards regression	C index 0.689 and 0.693. Distant metastasis free survival not different	Genetic data
Howard et al. (2020) [59]	HNC SCC	OS	33 527	DeepSurv, RSF *N*-MTLR	HR of 0.79, 0.83 and 0.90	Stage T4, HPV status, tonsil subsite
De Felice et al. (2020) [32]	OPC	OS	273	RF - Classification tree	Mean decrease accuracy of 4.29, 2.49 and 1.11 %	HPV status N status Early responders
Tran et al. (2020) [68]	HNC SCC	Local nodal response	32	LR, KNN, naive-Bayes	Accuracy 87.5 % with three feature model	Quantitative US radiomic features
Tosado et al. (2020) [70]	OPC	OS, RFS	644	RReliefF feature selector, Cox Model, RSF	AUC 0.6395 (OS) and 0.6483 (RFS)	Radiomic features combined with clinical features
Bogowicz et al. (2020) [71]	OPC, hypo pharynx, larynx, oral cavity	2YOS, HPV	1174	Feature selector (LR Z-Rad), Classification (hierarchical clustering, LR)	No significant differences in AUC between centralized and distributed.	981 radomic features
Rich et al. (2021) [69]	OPC HPV +	DM	225	Feature extractor (SMOTE, ADASYN, borderline SMOTE), SVM	AUC 0.84–0.95	CT based radiomic
Le et al. (2022) [72]	HNC SCC	10Y DM, Lr, OS	371	Cox Model, RF, CNN, DenseNet, InceptionV3, ResNet, ResNeXt and PreSANet	PreSAnet - Accuracy of 74 % (Lr) and 79 % (OS)	Performance decrease with PET images

Abbreviations: ADASYN, adaptive synthetic sampling; ANN, artificial neural network; BY, Bayesian; CT, Computed tomography; C-index, concordance index; CIFE, conditional informax feature extraction; CRT, chemo-radiotherapyDM, distant metastasis; DMFS, distant metastasis-free survival; HNC, head and neck cancer; HPV, human papillomavirus; HR, Hazard ratio; IAMB, incremental association Markov blanket; KNN, k nearest neighbor; LF, local failure; Lr, locoregional recurrence; LR, logistic regression; LRC, loco-regional control; NPC, nasopharyngeal cancer; MIFS, mutual information based feature selection; MRI, Magnetic resonance imaging; MRMR, Minimum redundancy feature selection; N, node; NN, nearest neighoor; *N*-MLTR, Neural Multi-Task Logistic Regression; OS, overall survival; PET, positron emission tomography; PCA, Principal component analysis; RF, random forest; RFC, relapse-free survival; RRF, Regularized random forest; RSF, random survival forest; SCC, squamous cell carcinoma; SMOTE, synthetic minority over-sampling technique; SVM, support vector machine; US, Ultrasounds; 2D CNN, two dimensions convolutional neural network; 3D CNN, three dimensions convolutional neural network.

locally advanced HNC treated with curative intent surgery, combined with adjuvant RT or CRT. Compared to the baseline model using clinicopathologic data alone, the identification and integration of genetic features associated with prognostic led to a model with better classification performance, with mean C-indexes of 0.689 (vs 0.673) and 0.693 (vs 0.678) for cancer-specific survival and locoregional recurrence-free survival, respectively. Also in the post-operative setting, Howard et al [59] aimed to build an overall survival (OS) predictive model to better identify patients that may benefit from the addition of adjuvant concurrent chemotherapy. The authors evaluated different ML models integrating a large retrospective data from the National Cancer Database, including 33 527 patients with squamous cell carcinoma of the oral cavity, oropharynx, hypopharynx, or larynx treated with definitive surgery followed by an adjuvant RT or CRT. Three different models were trained, then validated using a subset of the same cohort. Compared to RTOG 95–01 and EORTC 22931 recommendations [60–61], treatments guided by all three models had improved survival outcomes with hazard ratio (HR) of 0.79–0.90 and c index of 0.691–0.695 with similar accuracy and there was no survival benefit for CRT for patients recommended to receive RT alone. These models also identified important

variables related to prognosis such as year of diagnosis, T4, HPV positivity and tonsillar subsite.

In order to evaluate the role of radiomics for risk stratification of DM, Rich et al [69] used CT based radiomic and clinical features from 225 locally advanced OPC HPV+ patients treated with curative intent RT or CRT. Nine different algorithms were built using different radiomic datasets derived using different algorithms integrated to SVM classifier. All classifiers achieved at least an excellent level in discriminating the two patient cohorts with DM or not. In order to build a predictive model for OS and RFS outcomes, Tosado et al [70] incorporated retrospective CT based radiomic data with clinical data from 644 OPC patients treated with RT or CRT. Distinctive radiomic and clinical features with the cluster labels were identified and different supervised models were built using these features. Compared to the baseline model based on clinical features only, these models had better predictive performance for both outcomes. In another study by Jiang et al [62], aiming to predict outcomes in metastatic NPC patients treated with chemotherapy alone, RT alone or the combination of both, the retrospective hematological, clinical and therapeutic parameters of 347 patients were used in a SVM model. The multivariate model had a strong performance with an AUC at 0.761 and the classifier was able to stratify the patients into low risk and high risk groups with significantly different 2 year OS of 71.4 % vs 18.8 %, respectively. This classifier also helped identify that combined CRT was associated with significantly better outcomes in the low risk groups, but not in the higher risk groups.

Other studies focused on integrating radiomic data into ML algorithms to optimize predictive models for HNC treatment outcomes. In order to identify the optimal ML methods for radiomic-based overall survival prediction, Parmar et al [63] compared the performance of 13 features selection methods and 11 ML classification methods integrating CT based 440 radiomic features from 196 HNC patients. Three feature selection methods had the best performance with AUC between 0.66 and 0.69 and stability between 0.66 and 0.70 compared to the median values of AUC = 0.61 and stability = 0.66. Three classifying methods had the best performance with AUC between 0.61 and 0.67. Zdilar et al [64] used retrospective CT based radiomic, clinical, demographic, toxicity and cancer control outcome data from 529 OPC patients treated with curative intent RT or CRT to compare the predictive performance of different selectors for OS and RFS. Among 3800 radiomic features extracted, selected features using 8 different methods resulted in better AUC compared to clinical features alone. Among the feature selectors, RF based selectors had the best overall scores. In order to build a model predicting a radioresistance, Li et al [65] used retrospective data from 306 NPC patients treated with definitive CRT. Clinicopathological and radiomic/dosimetric features from planning CT and from follow up imagings including CT, MRI or PET. Once detected, recurrent tumor volumes were delineated, then was categorized as "in field recurrence" if the recurrence was inside the high-dose target. Eight discriminative features were identified from pretreatment MRIs compared between the patients with and without the disease recurrence. Features were fed to three different MLs, which were trained, then validated and yield accuracies ranging between 0.732 and 0.812. These results could indicate possible differences in heterogeneity in LR tumors. For a predictive performance of loco-regional recurrence (LRR), Starke et al [66] used the retrospective data from 291 patients with locally advanced HNC treated with CRT. A baseline Cox proportional hazards model (CPHM) using clinical features alone was compared to different 3D-CNN and 2D-CNN models built from scratch combining clinical features and CT images. Among these, the ensemble of 3D-CNNs had the best performance and successful validation with a C-index of 0.31. Patient risk group defined by this model's predictions showed significant differences in LRR with p = 0.001. The C-index for 2D-CNN and for CPHM was 0.38 and 0.39, respectively. In order to satisfy risks for DM and 5Y OS, Zhou et al [67] used MRI based radiomic with dosiomic and clinical features from 176 NPC patients treated with curative intent RT or CRT. With the radiomic features extracted, an algorithm model was built in order to classify into high- and low risk groups for DM. With the clinical features, the radiomic based models show strong predictive performance for both in the training and validation cohorts with AUC of 0.827 and 0.792 respectively. Another study was done to predict therapeutic response in metastatic lymph nodes by Tran et al [68]. Their team used the quantitative ultrasound based radiomic markers from 32 HNC patients with positive lymph nodes treated with curative intent RT or CRT. Depending on their 3 months follow up MRI, patients were divided into two different categories: complete responders or partial responders. Different radiomic features were extracted, then applied in LR, KNN and a naive-Bayes. Multi Parametric models showed a strong predictive power with high accuracy of 87.5 %. Significant differences in radiomic parameters were found between the two groups.

In order to evaluate the role of radiomics for risk stratification of DM, Rich et al [69] used CT based radiomic and clinical features from 225 locally advanced OPC HPV+ patients treated with curative intent RT or CRT. Nine different algorithms were built using different radiomic datasets derived using different algorithms integrated to SVM classifier. All classifiers achieved at least an excellent level in discriminating the two patient cohorts with DM or not. In order to build a predictive model for OS and RFS outcomes, Tosado et al [70] incorporated retrospective CT based radiomic data with clinical data from 644 OPC patients treated with RT or CRT. Distinctive radiomic and clinical features with the cluster labels were identified and different supervised models were built using these features. Compared to the baseline model based on clinical features only, these models had better predictive performance for both outcomes. In another study by Jiang et al [62], aiming to predict outcomes in metastatic NPC patients treated with chemotherapy alone, RT alone or the combination of both, the retrospective hematological, clinical and therapeutic parameters of 347 patients were used in a SVM model. The multivariate model had a strong performance with an AUC at 0.761 and the classifier was able to stratify the patients into low risk and high risk groups with significantly different 2 year OS of 71.4 % vs 18.8 %, respectively. This classifier also helped identify that combined CRT was associated with significantly better outcomes in the low risk groups, but not in the higher risk groups.

The sufficient quantity of a radomic data set is a common problem for an AI algorithm development and multicenter approach can be a solution but implicating ethical issues. In order to address this issue, Bogowicz et al [71] tested the distributed learning technique enabling training models on multicenter data without data leaving the hospitals. Two different approaches, centralized and distributed, were compared for 2Y OS and HPV status predictive models built with CT based radiomic, dosiomic and clinical features from 1174 HNC patients treated with curative intent RT or CRT. For both feature selection and classification, there was no significant difference in terms of performance between these two approaches. Most recently, Le et al [72] used retrospective cross-institutional patho-clinical and PET-CT based radiomic data set from 298 HNC patients treated with curative intent RT or CRT in order to train a predictive model based on a pseudo-volumetric convolutional neural network with PreSANet. The model was internally validated, then an extensive set of ablation experiments on the public data set showed AUROC of DM, LR and OS between 80 and 82 %. External validation on a retrospective dataset showed an AUROC at 69 % and a validation of single site-holdout and cross-validation showed mean accuracy across four different institutions was between 70 and 72 %.

Discussion

Recent studies have shown promising results in the use of ML in the field of HNC RT in predicting therapeutic outcomes and toxicity. Different algorithms have shown good predictive performance and have helped identify features that provide insight into the heterogeneous nature of HNC. Those features have included demographic characteristics, molecular, dosimetric, radiomic and therapeutic factors. Their integration to the current clinical decision algorithms has the potential

to improve risk stratification and selection of optimal therapeutic options.

Despite promising results, these models remain largely premature for clinical use at this stage. One major concern is the lack of standardization of the largely retrospective data used in a single center. In radiomics specifically, for each imaging modality, intra and inter-institutional variations in scanner, acquisition and reconstruction parameters have been shown to impact the robustness of the predictive models [73]. Pertaining to RT data, differences in tumor and organs at risk segmentations as well as dosimetric data further challenges reproducibility [42,74]. In addition, variations in feature extraction techniques, choice of robustness metrics and outcome definitions used across studies complicate the interpretation of results. The small sample sizes and high heterogeneity of HNC and frequent lack of external validation lead lack of generalizability of the current studies. Increased use of open-access data sets and multicenter prospective cohorts, along with strict guidelines for data standardization would be critical for these models to reach clinical usability [71–72,75]. In addition, training clinical oncologists to the field of AI, as well as integrating them early in the development and validation of these models will be increasingly important. As clinicians are ultimately responsible for decision making, ensuring their adequate understanding and interpretation of these models will help overcome the ''black box'' problem and facilitate clinical implementation of these tools [76–77].

Beside the risk stratification and decision support in HNC RT, AI has substantial potential in other upstream tasks such as automatic detection and segmentation of anatomical structures [13], automatic registration of mono-modal or multimodal images [14], temporal motion compensation [15], tumor and lesion grade classification [16], and nomograms for risk stratification-prognostic modeling [17]. AI can independently be introduced at all stages of a patient's treatment from diagnosis, to planning and re-planning, to long-term follow-up and prognosis, while allowing or benefiting from expert input along the way.

AI can reveal the radiomic signatures in HNC such as tumor characteristics including HPV status [29] or Programmed Death-Ligand 1 expression [30], identification of extranodal extension [31] as well as cancer control outcomes [32] and larynx/hypopharynx cancers [33]. Radiomics also has the potential to provide a quantitative assessment of tumor and normal tissue reaction to RT over the course of treatment (i.e. delta-radiomics) [34]. *The emerging use of multi-omics ((Gen-omics, Epigen-omics, Transcipt-omics, prote-omics, metabol-omics and microbi-omics)* [78–79] in oncology, along with the increasing quantity and quality of imaging in radiation oncology, represent a clear opportunity to propule the role of AI in HNC. *The integration* of daily imaging used over the course of RT in order to capture dynamically tumor response or early signs of toxicity could guide therapeutic decision or early interventions more efficiently, as suggested in early work using CBCTs data in dynamic predictive algorithms [80–81]. The increasing availability of the MR-Linac technology across institutions will lead to increased quality and both anatomic and functional information of daily RT imaging and will further increase the potential to unlock dynamic imaging-based biomarkers [82–85]. Finally, as the field of liquid biopsy if expanding rapidly, non invasive biologic serum or salivary HNC biomarkers [86–88] could be further integrated to these algorithms to increase the precision of dynamic clinical outcomes predictions over the course of treatment and in post-treatment follow-ups [89–94].

This review has several limitations. First, its non-systematic review method could cause potential risks of different bias with less transparency. However, this method seemed suitable to offer a general overview of the current stance of AI in HNC RT. Also, our study did not include a recently developed *radiomics quality score*, a tool built by Lambin et al. in order to determine the validity and completeness of radiomics studies [98]. For a more comprehensive understanding of the subject matter, future reviews could include use of this tool as well as an insight on the ongoing clinical trials evaluating AI and radiomics tools in HNC RT [99–101].

AI-models predicting cancel control and toxicity outcomes in HNC RT have shown promising performance and would be of high clinical utility for individualized risk-based decision making. These important challenges to the development and safe clinical implementation of these models could only be overcome with coordinated collaborative efforts to standardize, validate and expand these models to large enough datasets and test in the context of clinical trials.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Ferlay J, Sung H, Siegel RL, et al. Globalcancerstatistics 2020: GLOBOCANestimatesof incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71(3):209–49.

[2] Atun R, et al. Expanding global access to radiotherapy. Lancet Oncol 2015;16: 1153–86.

[3] Trotti A. Toxicity in head and neck cancer: a review of trends and issues. Int J Radiat Oncol Biol Phys 2000;47(1):1–12. https://doi.org/10.1016/s0360-3016(99)00558-1. PMID: 10758302.

[4] Fernandes Q, Merhi M, Raza A, Inchakalody VP, Abdelouahab N, Zar Gul AR, et al. Role of Epstein-barr virus in the pathogenesis of head and neck cancers and its potential as an immunotherapeutic target. Front Oncol 2018;6(8):257. https://doi.org/10.3389/fonc.2018.00257. PMID: 30035101; PMCID: PMC6043647.

[5] Di Credico G, Polesel J, Dal Maso L, et al. Alcohol drinking and head and neck cancer risk: the joint effect of intensity and duration. Br J Cancer 2020;123: 1456–63. https://doi.org/10.1038/s41416-020-01031-z.

[6] Caudell JJ, Torres-Roca JF, Gillies RJ, Enderling H, Kim S, Rishi A, et al. The future of personalised radiotherapy for head and neck cancer. Lancet Oncol 2017 May;18(5):e266–73. https://doi.org/10.1016/S1470-2045(17)30252-8. Epub 2017 Apr 26. PMID: 28456586; PMCID: PMC7771279.

[7] Alterio D, Marvaso G, Ferrari A, Volpe S, Orecchia R, Jereczek-Fossa BA. Modern radiotherapy for head and neck cancer. Semin Oncol 2019 Jun;46(3):233–45. https://doi.org/10.1053/j.seminoncol.2019.07.002. Epub 2019 Jul 26 PMID: 31378376.

[8] Grégoire V, Langendijk JA, Nuyts S. Advances in Radiotherapy for Head and Neck Cancer. J Clin Oncol 2015 Oct 10;33(29):3277–84. https://doi.org/10.1200/JCO.2015.61.2994. Epub 2015 Sep 8 PMID: 26351354.

[9] Giraud P, Giraud P, Gasnier A, et al. Radiomics and machine learning for radiotherapy in head and neck cancers. Front Oncol 2019;9:174.

[10] Parmar C, Grossmann P, Bussink J, et al. Machine learning methods for quantitative radiomic biomarkers. Sci Rep 2015;5:13087.

[11] Lambin P, Roelofs E, Reymen B, Velazquez ER, Buijsen J, Zegers CM, et al. Rapid Learning health care in oncology' - an approach towards decision support systems enabling customised radiotherapy. Radiother Oncol 2013 Oct;109(1):159–64. https://doi.org/10.1016/j.radonc.2013.07.007. Epub 2013 Aug 28 PMID: 23993399.

[12] Moor J. The dartmouth college artificial intelligence conference: the next fifty years. AI Mag 2006;27(4):87.

[13] Nikolov, S. et al. Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy. Preprint at arXiv (2018).

[14] Fan J, Cao X, Wang Q, Yap PT, Shen D. Adversarial learning for mono- or multi-modal registration. Med Image Anal. 2019 Dec;58:101545. 10.1016/j.media.2019.101545. Epub 2019 Aug 24. PMID: 31557633; PMCID: PMC7455790.

[15] Shi L, Han S, Zhao J, Kuang Z, Jing W, Cui Y, et al. Respiratory prediction based on multi-scale temporal convolutional network for tracking thoracic tumor movement. Front Oncol 2022 May;27(12):884523. https://doi.org/10.3389/fonc.2022.884523. PMID: 35692785; PMCID: PMC9184446.

[16] Reimagining T. Staging through artificial intelligence and machine learning image processing approaches in digital pathology Kaustav Bera, Ian Katz, and Anant Madabhushi. JCO Clin Cancer Inf 2020;4:1039–50.

[17] Machine Learning and Nomogram Prognostic Modeling for 2-Year Head and Neck Cancer–Specific Survival Using Electronic Health Record Data: A Multisite Study Damian P. Kotevski, Robert I. Smee, Claire M. Vajdic, and Matthew Field JCO Clinical Cancer Informatics 2023:7.

[18] Belgioia L, Morbelli SD, Corvò R. Prediction of response in head and neck tumor: focus on main hot topics in research. Front Oncol 2021 Jan;8(10):604965. https://doi.org/10.3389/fonc.2020.604965. PMID: 33489911; PMCID: PMC7821385.

[19] Niedzielski JS, Yang J, Stingo F, Liao Z, Gomez D, Mohan R, et al. A novel methodology using CT imaging biomarkers to quantify radiation sensitivity in the esophagus with application to clinical trials. Sci Rep 2017;7:6034. https://doi.org/10.1038/s41598-017-05003-x.

[20] Carvalho S, Troost EGC, Bons J, Menheere P, Lambin P, Oberije C. Prognostic value of blood-biomarkers related to hypoxia, inflammation, immune response

and tumour load in non-small cell lung cancer – a survival model with external validation. Radiother Oncol 2016;119:487–94.
[21] Gonçalves M, Gsaxner C, Ferreira A, Li J, Puladi B, Kleesiek J, et al. Radiomics in head and neck cancer outcome predictions. Diagnostics 2022;12:2733.
[22] Koçak B, Durmaz EŞ, Ateş E, Kılıçkesmez Ö. Radiomics with artificial intelligence: a practical guide for beginners. Diagn Interv Radiol 2019 Nov;25(6):485–95. https://doi.org/10.5152/dir.2019.19321. PMID: 31650960; PMCID: PMC6837295.
[23] Nardone V, Reginelli A, Grassi R, et al. Delta radiomics: a systematic review. Radiol med 2021;126:1571–83.
[24] Newbold K, Powell C. PET/CT in radiotherapy planning for head and neck cancer. Front Oncol 2012 Dec;10(2):189. https://doi.org/10.3389/fonc.2012.00189. PMID: 23233906; PMCID: PMC3518254.
[25] Eyassu E, Young M. Nuclear Medicine PET/CT Head and Neck Cancer Assessment, Protocols, And Interpretation. 2022 Nov 28. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan–. PMID: 34424632.
[26] Yan D, Lockman D, Martinez A, Wong J, Brabbins D, Vicini F, et al. Computed tomography guided management of interfractional patient variation. Semin Radiat Oncol 2005;15:168–79. https://doi.org/10.1016/j.semradonc.2005.01.007.
[27] Mutic S, Dempsey JF. The ViewRay system: magnetic resonance-guided and controlled radiotherapy. Semin Radiat Oncol 2014;24:196–9. https://doi.org/10.1016/j.semradonc.2014.02.008.
[28] Simo R, Homer J, Clarke P, Mackenzie K, Paleri V, Pracy P, et al. Follow-up after treatment for head and neck cancer: United Kingdom National Multidisciplinary Guidelines. J Laryngol Otol 2016 May;130(S2):S208–11.
[29] Sohn B, Choi YS, Ahn SS, Kim H, Han K, Lee SK, et al. Machine learning based radiomic HPV phenotyping of oropharyngeal SCC: A feasibility study using MRI. Laryngoscope 2021 Mar;131(3):E851–6. https://doi.org/10.1002/lary.28889. Epub 2020 Jul 13 PMID: 33070337.
[30] Ren Q, Xiong F, Zhu P, Chang X, Wang G, He N, et al. Assessing the robustness of radiomics/deep learning approach in the identification of efficacy of anti-PD-1 treatment in advanced or metastatic non-small cell lung carcinoma patients. Front Oncol 2022 Aug;5(12):952749. https://doi.org/10.3389/fonc.2022.952749. PMID: 35992867; PMCID: PMC9390967.
[31] Kann BH, Aneja S, Loganadane GV, Kelly JR, Smith SM, Decker RH, et al. Pretreatment identification of head and neck cancer nodal metastasis and extranodal extension using deep learning neural networks. Sci Rep 2018;8:14036. https://doi.org/10.1038/s41598-018-32441-y.
[32] De Felice F, Humbert-Vidan L, Lei M, King A, Guerrero UT. Analyzing oropharyngeal cancer survival outcomes: a decision tree approach. Br J Radiol 2020;93(1111):20190464.
[33] Zhong J, Frood R, Brown P, Nelstrop H, Prestwich R, McDermott G, et al. Machine learning-based FDG PET-CT radiomics for outcome prediction in larynx and hypopharynx squamous cell carcinoma. Clin Radiol 2021 Jan;76(1):78.e9–78.e17. https://doi.org/10.1016/j.crad.2020.08.030. Epub 2020 Oct 6 PMID: 33036778.
[34] Fatima K, Dasgupta A, DiCenzo D, Kolios C, Quiaoit K, Saifuddin M, et al. Ultrasound delta-radiomics during radiotherapy to predict recurrence in patients with head and neck squamous cell carcinoma. Clin Transl Radiat Oncol 2021 Mar;12(28):62–70. https://doi.org/10.1016/j.ctro.2021.03.002. PMID: 33778174; PMCID: PMC7985224.
[35] Kermany DS, Goldbaum M, Cai W, Valentim CC, Liang H, Baxter SL, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 2018;172(5):1122–31.
[36] Yang R, Yu Y. Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis. Front Oncol 2021 Mar;9(11):638182. https://doi.org/10.3389/fonc.2021.638182. PMID: 33768000; PMCID: PMC7986719.
[37] Deist TM, Dankers FJ, Valdes G, Wijsman R, Hsu I, Oberije C, et al. Erratum: Machine learning algorithms for outcome prediction in (chemo)radiotherapy: An empirical comparison of classifiers. Med Phys 2019;46:1080–7.
[38] Christianen MEMC, van der Schaaf A, van der Laan HP, Verdonck-de Leeuw IM, Doornaert P, Chouvalova O, et al. Swallowing sparing intensity modulated radiotherapy (SW-IMRT) in head and neck cancer: clinical validation according to the model-based approach. Radiother Oncol 2016;118:298–303.
[39] Houweling AC, Philippens MEP, Dijkema T, Roesink JM, Terhaard CHJ, Schilstra C, et al. A comparison of dose-response models for the parotid gland in a large group of head-and-neck cancer patients. Int J Radiat Oncol Biol Phys 2010;76(4):1259–65.
[40] Buettner F, Miah AB, Gulliford SL, Hall E, Harrington KJ, Webb S, et al. Novel approaches to improve the therapeutic index of head and neck radiotherapy: an analysis of data from the PARSPORT randomised phase III trial. Radiother Oncol 2012;103(1):82–7.
[41] Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of normal tissue complication probability models in the clinic. Int J Radiat Oncol Biol Phys 2010;76:10–9.
[42] Christianen ME, Langendijk JA, Westerlaan HE, van de Water TA, Bijl HP. Delineation of organs at risk involved in swallowing for radiotherapy treatment planning. Radiother Oncol 2011;101(3):394–402.
[43] El Naqa I, Bradley JD, Lindsay PE, Hope AJ, Deasy JO. Predicting radiotherapy outcomes using statistical learning techniques. Phys Med Biol 2009;54(18):S9–.
[44] Gabrys HS, Buettner F, Sterzing F, Hauswald H, Bangert M. Design and selection of machine learning methods using radiomics and dosiomics for normal tissue complication probability modeling of xerostomia. Front Oncol 2018.
[45] Jiang W, Lakshminarayanan P, Hui X, Han P, Cheng Z, Bowers M, et al. Machine learning methods uncover radiomorphologic dose patterns in salivary glands that predict xerostomia in patients with head and neck cancer. Adv Radiat Oncol 2019;4:401–12.
[46] Dijk L, Brouwer CL, Schaaf AV, Burgerhof JG, Beukinga RJ, Langendijk JA, et al. CT image biomarkers to improve patient-specific prediction of radiation-induced xerostomia and sticky saliva. Radiother Oncol 2017;122:185–91.
[47] Beasley W, Thor M, McWilliam A, Green A, Mackay R, Slevin N, et al. Image-based data mining to probe dosimetric correlates of radiation-induced trismus. Int J Radiat Oncol Biol Phys 2018 November 15;102(4):1330–8.
[48] Terrell JE, Ronis DL, Fowler KE, Bradford CR, Chepeha DB, Prince ME, et al. Clinical predictors of quality of life in patients with head and neck cancer. Arch Otolaryngol Head Neck Surg 2004;130:401–8.
[49] Dean J, Wong K, Gay H, Welsh L, Jones AB, Schick U, et al. Incorporating spatial dose metrics in machine learning-based normal tissue complication probability (NTCP) models of severe acute dysphagia resulting from head and neck radiotherapy. Clin Transl Radiat Oncol 2018;8:27–39.
[50] Reddy J, Lindsay W, Berlind C, Ahern C, Holmes A, Smith B, et al. Applying a machine learning approach to predict acute radiation toxicities for head and neck cancer patients. Int J Radiat Oncol Biol Phys 2019;105:S69.
[51] Wojcieszynski A, Cava WL, Baumann B, Lukens J, Ghiam AF, Urbanowicz R, et al. Machine learning to predict toxicity in head and neck cancer patients treated with definitive chemoradiation. Int J Radiat Oncol Biol Phys 2019;105:E139–40.
[52] Cheng Z, Nakatsugawa M, Hu C, Robertson SP, Hui X, Moore JA, et al. Evaluation of classification and regression tree (CART) model in weight loss prediction following head and neck cancer radiation therapy. Adv Radiat Oncol 2017 Dec 7;3(3):346–55.
[53] Dean JA, Wong KH, Welsh LC, Jones A, Schick U, Newbold KL, et al. Normal tissue complication probability (NTCP) modelling using spatial dose metrics and machine learning methods for severe acute oral mucositis resulting from head and neck radiotherapy. Radiother Oncol 2016;120:21–7. https://doi.org/10.1016/j.radonc.2016.05.015.
[54] Zhang B, Lian Z, Zhong L, et al. Machine-learning based MRI radiomics models for early detection of radiation-induced brain injury in nasopharyngeal carcinoma. BMC Cancer 2020;20:502.
[55] Laia Humbert-Vidan, Vinod Patel, Robin Andlauer, Andrew P King, and Teresa Guerrero Urbano. 2022. Prediction of Mandibular ORN Incidence from 3D Radiation Dose Distribution Maps Using Deep Learning. In Applications of Medical Artificial Intelligence: First International Workshop, AMAI 2022, Held in Conjunction with MICCAI 2022, Singapore, September 18, 2022, Proceedings. Springer-Verlag, Berlin, Heidelberg, 49–58.
[56] Brierley J, Gospodarowicz M, Wittekind C. UICC TNM classification of malignant tumours. eighth ed. Chichester: Wiley; 2017.
[57] Kaka AS, Kumar B, Kumar P, Wakely PE Jr, Kirsch CM, Old MO, Ozer E, Agrawal A, Carrau RE, Schuller DE, Siddiqui F, Teknos TN. Highly aggressive human papillomavirus-related oropharyngeal cancer: clinical, radiologic, and pathologic characteristics. Oral Surg Oral Med Oral Pathol Oral Radiol. 2013 Sep;116(3):327-35. 10.1016/j.oooo.2013.04.011. Epub 2013 Jun 14. PMID: 23770280; PMCID: PMC3748144.
[58] Tseng Y-J, Wang H-Y, Lin T-W, Lu J-J, Hsieh C-H, Liao C-T. Development of a machine learning model for survival risk stratification of patients with advanced oral cancer. JAMA Netw Open 2020;3(8):e2011768.
[59] Howard FM, Kochanny S, Koshy M, Spiotto M, Pearson AT. Machine learning-guided adjuvant treatment of head and neck cancer. JAMA Netw Open 2020;3(11):e2025881.
[60] Bernier J, Domenge C, Ozsahin M, etal; European Organization for Research and Treatment of Cancer Trial 22931. Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. N Engl J Med. 2004;350(19):1945-1952.
[61] Cooper JS, Pajak TF, Forastiere AA, et al. Radiation Therapy Oncology Group 9501/Intergroup. Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck. N Engl J Med 2004;350(19):1937–44.
[62] Jiang R, You R, Pei XQ, Zou X, Zhang MX, Wang TM, et al. Development of a ten-signature classifier using a support vector machine integrated approach to subdivide the M1 Stage Into M1a and M1b stages of nasopharyngeal carcinoma with synchronous metastases to better predict patients' survival. Oncotarget 2016;7(3):3645–57.
[63] Parmar C, Grossmann P, Rietveld D, Rietbergen MM, Lambin P, Aerts HJWL. Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer. Front Oncol 2015;5:272.
[64] Zdilar L, Vock DM, Marai GE, Fuller CD, Mohamed ASR, Elhalawani H, et al. Evaluating the Effect of Right-Censored End Point Transformation for Radiomic Feature Selection of Data From Patients With Oropharyngeal Cancer. JCO Clin Cancer Inform (2018) 2):1–19.
[65] Li S, Wang K, Hou Z, Yang J, Ren W, Gao Z, et al. Use of radiomics combined with machine learning method in the recurrence patterns after intensity-modulated radiotherapy for nasopharyngeal carcinoma: a preliminary study. Front Oncol 2018;8:648.
[66] Starke S, Leger S, Zwanenburg A, Leger K, Lohaus F, Linge A, et al. 2D and 3D convolutional neural networks for outcome modelling of locally advanced head and neck squamous cell carcinoma. Sci Rep 2020;10(1):15625.
[67] Zhou Z, Wang K, Folkert M, et al. Multifaceted radiomics for distant metastasis prediction in head & neck cancer. Phys Med Biol 2020;65(15):155009.

[68] Tran WT, Suraweera H, Quaioit K, Cardenas D, Leong KX, Karam I, et al. Predictive quantitative ultrasound radiomic markers associated with treatment response in head and neck cancer. Future Sci OA 2020;6(1). FSO433.
[69] Rich B, Huang J, Yang Y, Jin W, Johnson P, Wang L, et al. Radiomics predicts for distant metastasis in locally advanced human papillomavirus-positive oropharyngeal squamous cell carcinoma. Cancers (Basel) 2021;13(22):5689. https://doi.org/10.3390/cancers13225689. PMID: 34830844; PMCID: PMC8616361.
[70] Tosado J, Zdilar L, Elhalawani H, et al. Clustering of largely right-censored oropharyngeal head and neck cancer patients for discriminative groupings to improve outcome prediction. Sci Rep 2020;10(1):3811.
[71] Bogowicz M, Jochems A, Deist TM, et al. Privacy-preserving distributed learning of radiomics to predict overall survival and HpV status in head and neck cancer. Sci Rep 2020;10(1):4542.
[72] Le WT, Vorontsov E, Romero FP, et al. Cross-institutional outcome prediction for head and neck cancer patients using self-attention neural networks. Sci Rep 2022; 12:3183.
[73] Reiazi R, Abbas E, Famiyeh P, Rezaie A, Kwan JYY, Patel T, et al. The impact of the variation of imaging parameters on the robustness of Computed Tomography radiomic features: A review. Comput Biol Med 2021 Jun;133:104400. https://doi.org/10.1016/j.compbiomed.2021.104400. Epub 2021 Apr 16 PMID: 33930766.
[74] Larue RTHM, van Timmeren JE, de Jong EEC, Feliciani G, Leijenaar RTH, Schreurs WMJ, et al. Influence of gray level discretization on radiomic feature stability for different CT scanners, tube currents and slice thicknesses: a comprehensive phantom study. Acta Oncol 2017 Nov;56(11):1544–53. https://doi.org/10.1080/0284186X.2017.1351624. Epub 2017 Sep 8 PMID: 28885084.
[75] Zwanenburg A, et al. The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high throughput image-based phenotyping. Radiology 2020.
[76] Gilpin, L. H. et al. in 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA) 80–89 (2018).
[77] Artificial Intelligence in Oncology: Current Capabilities, Future Opportunities, and Ethical Considerations Jacob T. Shreve, MD, Sadia A. Khanani, MD, and Tufia C. Haddad, MD American Society of Clinical Oncology Educational Book 2022 :42, 842-851.
[78] Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. Genome Biol 2017; 18:83.
[79] Yan J, Risacher SL, Shen L, Saykin AJ. Network approaches to systems biology analysis of complex disease: integrative methods for multi-omics data. Brief Bioinform 2018;19:1370–81.
[80] Sumner W, Kim SS, Vitzthum L, et al. End of treatment cone-beam computed tomography (CBCT) is predictive of radiation response and overall survival in oropharyngeal squamous cell carcinoma. Radiat Oncol 2021;16:147.
[81] Muelas-Soria R, García-Mollá R, Morillo-Macías V, Bonaque-Alandí J, Sorribes-Carreras P, García-Piñón F, et al. The usefulness of adaptative radiotherapy in prostate cancer: how, when, and who? Biomedicines 2022;10:1401.
[82] Chin S, Eccles CL, McWilliam A, Chuter R, Walker E, Whitehurst P, et al. Magnetic resonance-guided radiation therapy: A review. J Med Imaging Radiat Oncol 2020; 64:163–77. https://doi.org/10.1111/1754-9485.12968.
[83] Jager EA, Ligtenberg H, Caldas-Magalhaes J, Schakel T, Philippens ME, Pameijer FA, et al. Validated guidelines for tumor delineation on magnetic resonance imaging for laryngeal and hypopharyngeal cancer. Acta Oncol Stockh Swed 2016;55:1305–12. https://doi.org/10.1080/0284186X.2016.1219048.
[84] King AD, Thoeny HC. Functional MRI for the prediction of treatment response in head and neck squamous cell carcinoma: potential and limitations. Cancer Imaging 2016;16:23. https://doi.org/10.1186/s40644-016-0080-6.
[85] Watson GA, Taylor K, Siu LL. Innovation and advances in precision medicine in head and neck cancer. Crit Issues Head Neck Oncol 2021;24:355–73. https://doi.org/10.1007/978-3-030-63234-2_24.
[86] Shanmugam A, Hariharan AK, Hasina R, Nair JR, Katragadda S, Irusappan S, et al. Ultrasensitive detection of tumor-specific mutations in saliva of patients with oral cavity squamous cell carcinoma. Cancer 2021.
[87] Arantes LMRB, De Carvalho AC, Melendez ME, Carvalho AL. Serum, plasma and saliva biomarkers for head and neck cancer. Expert Rev Mol Diagn 2018;18: 85–112.
[88] Kong L, Birkeland AC. Liquid biopsies in head and neck cancer: current state and future challenges. Cancers 1874;2021:13.
[89] Jatana KR, Balasubramanian P, Lang JC, Yang L, Jatana CA, White E, et al. Significance of circulating tumor cells in patients with squamous cell carcinoma of the head and neck: Initial results. Arch Otolaryngol Head Neck Surg 2010;136: 1274–9.
[90] Inhestern J, Oertel K, Stemmann V, Schmalenberg H, Dietz A, Rotter N, et al. Prognostic role of circulating tumor cells during induction chemotherapy followed by curative surgery combined with postoperative radiotherapy in patients with locally advanced oral and oropharyngeal squamous cell cancer. PLoS One 2015;10:e0132901.
[91] Strati A, Koutsodontis G, Papaxoinis G, Angelidis I, Zavridou M, Economopoulou P, et al. Prognostic significance of PD-L1 expression on circulating tumor cells in patients with head and neck squamous cell carcinoma. Ann Oncol Off J Eur Soc Med Oncol 2017;28:1923–33.
[92] Chera BS, Kumar S, Shen C, Amdur R, Dagan R, Green R, et al. Plasma circulating tumor HPV DNA for the surveillance of cancer recurrence in HPV-associated oropharyngeal cancer. J Clin Oncol Off J Am Soc Clin Oncol 2020;38:1050–8.
[93] Rutkowski TW, Mazurek AM, Snietura M, Hejduk B, Jedrzejewska M, Bobek-Billewicz B, D'Amico A, Pigłowski W, Wygoda A, Składowski K, et al. Circulating HPV16 DNA may complement imaging assessment of early treatment efficacy in patients with HPV-positive oropharyngeal cancer. J Transl Med 2020;18:1–10.
[94] Theodoraki M-N, Yerneni SS, Brunner C, Theodorakis J, Hoffmann TK, Whiteside TL. Plasma-derived exosomes reverse epithelial-to-mesenchymal transition after photodynamic therapy of patients with head and neck cancer. Oncoscience 2018;5:75–87.
[95] Pota M, Scalco E, Sanguineti G, Farneti A, Cattaneo GM, Rizzo G, et al. Early prediction of radiotherapy-induced parotid shrinkage and toxicity based on CT radiomics and fuzzy classification. Artif Intell Med 2017;81:41–53. https://doi.org/10.1016/j.artmed.2017.03.004.
[96] Fujima N, Shimizu Y, Yoshida D, Kano S, Mizumachi T, Homma A, et al. Machine-learning-based prediction of treatment outcomes using MR imaging-derived quantitative tumor information in patients with sinonasal squamous cell carcinomas: a preliminary study. Cancers 2019;11(6):800.
[97] Wu J, Gensheimer MF, Zhang N, Han F, Liang R, Qian Y, et al. Integrating tumor and nodal imaging characteristics at baseline and mid-treatment computed tomography scans to predict distant metastasis in oropharyngeal cancer treated with concurrent chemoradiotherapy. Int J Radiat Oncol 2019;104(4):942–52. https://doi.org/10.1016/j.ijrobp.2019.03.036.
[98] Lambin P, Leijenaar R, Deist T, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749–62.
[99] Ciro Franzese, Federico Fornasier. AI for Head Neck Cancer Treated With Adaptive RadioTherapy (RadiomicART) (RadiomicArt). NCT05081531.
[100] Ye Zhang, Li Ma. Deep Learning-based Classification and Prediction of Radiation Dermatitis in Head and Neck Patients. NCT05607225.
[101] Panagiotis Balermpas. MRI – Guided Adaptive RadioTHerapy for Reducing XerostomiA in Head and Neck Cancer (MARTHA-trial) (MARTHA). NCT03972072.