International Journal of Medical Informatics 188 (2024) 105464

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

Artificial Intelligence-Driven Radiomics in Head and Neck Cancer: Current Status and Future Prospects

Rasheed Omobolaji Alabi [a,b,*], Mohammed Elmusrati [b], Ilmo Leivo [c], Alhadi Almangush [1,a,c,d,e], Antti A. Mäkitie [1,a,f,g]

[a] Research Program in Systems Oncology, Faculty of Medicine, University of Helsinki, Helsinki, Finland
[b] Department of Industrial Digitalization, School of Technology and Innovations, University of Vaasa, Vaasa, Finland
[c] University of Turku, Institute of Biomedicine, Pathology, Turku, Finland
[d] Department of Pathology, University of Helsinki, Helsinki, Finland
[e] Faculty of Dentistry, Misurata University, Misurata, Libya
[f] Department of Otorhinolaryngology – Head and Neck Surgery, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[g] Division of Ear, Nose and Throat Diseases, Department of Clinical Sciences, Intervention and Technology, Karolinska Institute and Karolinska University Hospital, Stockholm, Sweden

A R T I C L E I N F O

Keywords:
Radiomics
Artificial Intelligence
Deep Learning
Machine learning
Head and Neck Cancer
Systematic Review

A B S T R A C T

Background: Radiomics is a rapidly growing field used to leverage medical radiological images by extracting quantitative features. These are supposed to characterize a patient's phenotype, and when combined with artificial intelligence techniques, to improve the accuracy of diagnostic models and clinical outcome prediction.
Objectives: This review aims at examining the application areas of artificial intelligence-based radiomics (AI-based radiomics) for the management of head and neck cancer (HNC). It further explores the workflow of AI-based radiomics for personalized and precision oncology in HNC. Finally, it examines the current challenges of AI-based radiomics in daily clinical oncology and offers possible solutions to these challenges.
Methods: Comprehensive electronic databases (PubMed, Medline via Ovid, Scopus, Web of Science, CINAHL, and Cochrane Library) were searched following the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines. The quality of included studies and their risk of biases were evaluated using the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) and Prediction Model Risk of Bias Assessment Tool (PROBAST).
Results: Out of the 659 search hits retrieved, 45 fulfilled the inclusion criteria. Our review revealed that the application of AI-based radiomics model as an ancillary tool for improved decision-making in HNC management includes radiomics-based cancer diagnosis and radiomics-based cancer prognosis. The radiomics-based cancer diagnosis includes tumor staging, tumor grading, and classification of malignant and benign tumors. Similarly, radiomics-based cancer prognosis includes prediction for treatment response, recurrence, metastasis, and survival. In addition, the challenges in the implementation of these models for clinical evaluations include data imbalance, feature engineering (extraction and selection), model generalizability, multi-modal fusion, and model interpretability.
Conclusion: Considering the highly subjective and interobserver variability that is peculiar to the interpretation of medical images by expert clinicians, AI-based radiomics seeks to offer potentially useful quantitative information, which is not visible to the human eye or unintentionally often remain ignored during clinical imaging practice. By enabling the extraction of this type of information, AI-based radiomics has the potential to revolutionize HNC oncology, providing a platform for more personalized, higher quality, and cost-effective care for HNC patients.

* Corresponding author at: Research Program in Systems Oncology, Faculty of Medicine, University of Helsinki, Helsinki, Finland.
E-mail address: rasheed.alabi@helsinki.fi (R.O. Alabi).
[1] The last two authors have equal contributions.

https://doi.org/10.1016/j.ijmedinf.2024.105464
Received 15 October 2023; Received in revised form 20 April 2024; Accepted 22 April 2024
Available online 23 April 2024

1. Introduction

Head and neck cancer (HNC) accounts for over 660,000 new cases and 325,000 associated deaths annually, and as a result, it represents the seventh most common cancer worldwide [1–3]. Approximately 90% of HNCs are squamous cell carcinomas [3]. The major general risks are tobacco smoking and/or alcohol consumption [3], and human papilloma virus (HPV) infection for oropharyngeal cancer [4,5]. The mortality ratio for HNC is high due to factors such as late (advanced stage) presentation relating to its characteristic insidious and asymptomatic development and absence of methods for early detection [6]. There have been improvements in treatment protocols including surgery and (chemo)radiotherapy either alone or in combination [7–9]. However, the therapeutic approaches are challenging and cause changes in breathing, swallowing, speech, and these treatment-related adverse effects affect management outcome of HNC. Consequently, HNC has been characterized by a high ratio of decreased quality of life and survival [6].

Outcome prognostication has been marked as one of the approaches to improve the management of HNC [10,11]. Recently, myriads of published articles have focused on the application of subfields of artificial intelligence (AI) such as machine learning (ML) and deep learning (DL) for models that may improve clinical decision-making processes [12–15]. Additionally, these models have ensured automated evaluation of medical images and can potentially enhance the workflow management of radiation oncology [6,7]. While the ML approach relies on tabular data for model development, the DL approach has been primarily focused on medical images. However, recent advancements in medical imaging have shown that large amounts of additional quantitative information, otherwise undetectable, can be extracted from medical images of HNC patients [6]. This marked the emergence of radiomics as an approach for extracting some high-dimensional data from clinical imaging data [6,7].

Radiomics is a fast and cost-efficient procedure for extracting quantitative features from clinical images to provide specific information on tumor heterogeneity, texture, intensity, and morphology information [7]. This information could further assist the clinicians to infer data on tumor histology, grade, metabolism, recurrence, distant metastasis, and even patient survival [7]. Considering the fact that magnetic resonance imaging (MRI) is the modality of choice for diagnosis, staging, and evaluation of response to treatment in HNC [6], radiomics is posited to generate reliable prognostic and biological information from image data. Besides MRI, computed tomography (CT) and spectral photon-counting CT have also been used in the diagnostics and prognostics of HNC [16]. Radiomics approach on clinical images (MRI or CT) combined with ML/DL techniques can provide precise stratification of HNC patients, providing more parameters to aid individualized diagnosis and treatment [17].

Several studies emphasizing the potential of radiomics and AI (ML and DL) techniques have been published in recent years [6]. Remarkably, the application of this emerging field of radiomics and DL/ML techniques in HNC is gaining significant attention [6]. Therefore, this study primarily aims to systematically review the published studies on the application of AI-based radiomics in HNC to identify the specific areas of its application, clinical concerns and limitations, and the potential for future studies. Furthermore, we aim to summarize the basic workflow of AI-based radiomics in HNC.

2. Materials and methods

2.1. Search of databases

The entire process of database search, screening, inclusion and exclusion, and reporting of the potentially relevant studies followed the Preferred Reporting Items for Systematic Review and Meta-Analysis

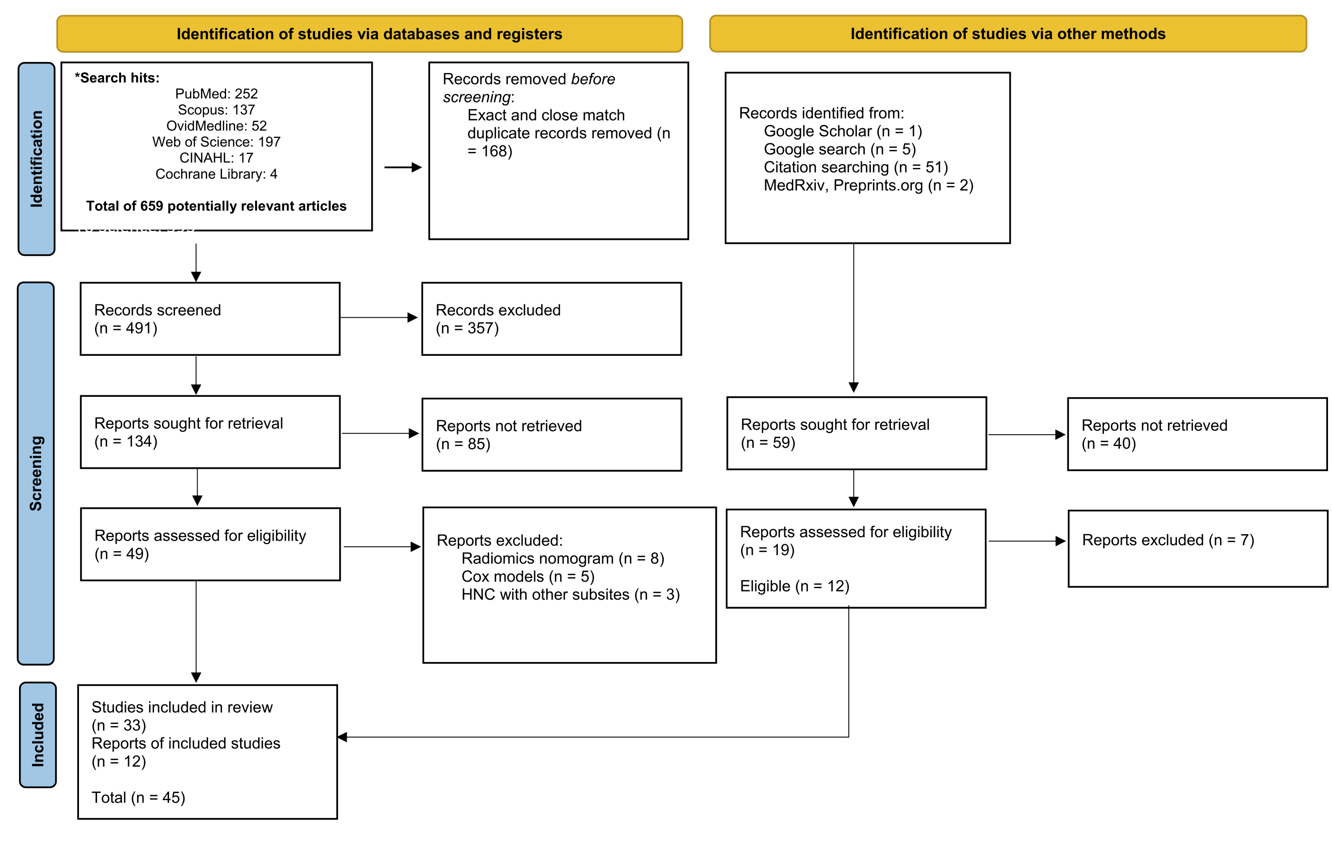

Fig. 1. The PRISMA flow chart.

(PRISMA) [Fig. 1]. A total of five (6) databases – Medline via Ovid, PubMed, Scopus, Web of Science, CINAHL, and Cochrane Library were systematically searched (Fig. 1).

2.2. Search terms and period

The potentially relevant articles were retrieved by combining search keywords: [('Artificial Intelligence OR Machine Learning') AND ('head and neck cancer') AND ('Radiomics')] from these databases from inception until 31st of August 2023 (Fig. 1). Additionally, Google Scholar and the reference lists of potentially relevant articles were searched to reduce research waste and maximize grey literatures.

2.3. Inclusion and exclusion criteria

All studies that combined radiomics with any of the subfields of AI such as ML or DL for diagnosis or prognosis in HNC were included. The potential studies to be included were further analyzed based on the PICO model – Population (HNC patients), Intervention (radiomics + ML), Comparison (results presentation), and Outcome (diagnosis or prognosis) prior to inclusion in Table 1. Comments, opinions, perspectives, guidelines, case studies, editorials, reviews, and articles in languages other than English were excluded. All studies that specifically examined the statistical approaches only or the Cox proportional hazards regression model for radiomics were excluded. This is necessary to homogenize the result of this study. In addition, considering the objective of this study which was centered around analyzing the performance of machine and deep learning algorithms on radiomics-generated features, excluding statistical approach was warranted in order to different between methodological approaches. Similarly, all studies that focused on radiomics-based nomograms were excluded. Studies that focused on cancer subsites other than HNC were also excluded.

2.4. Search analysis and screening

All the retrieved potentially relevant articles from the database search were exported to Endnote where duplicates and irrelevant studies were removed. Following the removal of duplicates and irrelevant studies, two independent researchers (R.A & A.A) further performed screening of potentially relevant articles. This screening was done in two distinct but successive phases. In the first phase, the titles and abstracts were analyzed in relation to the research objectives of this study. In the second phase, a comprehensive full-text assessment of the potential articles was performed for relevance to the objectives of this study.

2.5. Reporting of eligible studies

The same two independent researchers discussed how to resolve possible discrepancies in articles to be included in this study. A third-party (M.A) adjudication or consensus meeting was organized for disagreements and discussion where necessary. The inter-observer reliability between these researchers was measured using Cohen's Kappa coefficient ($k = 0.93$). A data extraction sheet was used to summarize the content of eligible studies based on the agreement between the two independent researchers. All eligible studies included in this review were summarized in Table 2.

2.6. Data extraction

For each eligible article (Table 2), article information (the first author's name, title of article, year of publication, country), article details (number of patients, imaging techniques, subsite, radiomics approach), article results (analyzed endpoints and performance/statistical findings), and article conclusions were retrieved and reported in Table 2. The workflow of AI-based radiomics in each article in Table 2 were noted. This workflow is summarized in Fig. 2. The challenges, limitations, and clinical concerns mentioned in these studies were noted (Fig. 3).

2.7. Quality appraisal

The quality appraisal of the included studies was done using the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) (Supplementary I) [18–21].

2.8. Risk of bias analysis

The risk of bias (RoB) of the developed prediction model for each eligible study was assessed using the risk of bias assessment tool (PROBAST). PROBAST assesses the quality and applicability of included studies [22,23] (Table 3). The PROBAST assessment covers 4 domain areas – participants, predictors, outcome, and analysis. Each domain has a set of questions to assist in assessing the domain. The answers to these questions have been predefined as *yes (+)*, *no (−)*, or *no information/unclear (?)*. The interpretation of these answers is that – *yes* (shows no/low bias), *no* indicates the possibility of bias (high bias) and unclear means unclear concerning applicability and RoB. The overall risk of bias was low if the four domains were graded low. The total risk of bias was high if at least 1 domain was scored high. The overall risk of bias was identified as unclear if the risk of bias was unclear for at least one domain area and low for the other domain areas. The details of the bias analysis and the corresponding results from each examined bias are given in Table 3.

2.9. Research Ethics Committee approval

Research Ethics Committee approval was not applicable to this review. The analysed articles in this review were based on previously conducted studies (Table 2). Therefore, this review does not contain any new studies with human participants or animals performed by any of the authors.

Table 1
Inclusion and exclusion using modified PICO model.

Selection criteria	Inclusion criteria	Exclusion criteria
P: population	Studies that examined application of machine learning (ML) or deep learning (DL) on radiomics generated features from computed tomography (CT), positron emission tomography (PET), magnetic resonance imaging (MRI), and ultrasound images for diagnosis or prognosis in head and neck cancer (HNC)	Other sites apart from HNC. General overview, editorial, narrative studies, systematic review, and comments
I: intervention	ML or DL-based radiomics model for diagnosis or prognosis in HNC	Traditional statistical method, Cox proportional model, and radiomics nomogram
C: comparison	Model performance based on any of the widely used ML metrics.	No performance metrics reported
O: outcome	Examined specific endpoints either diagnosis, prognosis, or both.	General overview, editorial, narrative studies, and comments

Table 2
Extracts of the main findings from the included studies.

Studies									
Author, year of publication (Country)	**Head and Neck or its subsite**	**Application (end point)**	**Data modality (source)**	**No of patients**	**Dataset model training**	**Software used for radiomics**	**Radiomics-based ML/DL approach**	**Model evaluation**	**Conclusion**
Folkert et al., 2017 (USA)	OPC	Outcome prediction (OS)	FDG-PET	174	Leave-one-out cross validation	MATLAB	LR	For ACM: AUC: 0.65 For LF: AUC: 0.73 For DM: AUC: 0.66	Radiomics-based models could assist in patient selection for dose reduction following identification of low-risk patients for local failure, or treatment intensification with additional adjuvant chemotherapy following identification of high-risk patients for distant failure
Vallieres et al., 2017 (Canada)	HNC	Outcome prediction (LR, DM, & OS)	FDG-PET/CT (Local)	300	194 training; 106 testing	MATLAB	RF	For locoregional recurrence prediction: AUC = 0.69. For distant metastases: AUC: 0.86	Radiomics could allow for a better personalization of chemo-radiation treatments for HNC patients from different risk groups.
Wang et al., 2017 (China & USA)	Stage II-IV NPC	Induction control response	(2D) T1WI, T2WI, T2WI FS, and CE T1WI	120	–	R software	LASSO & LR	AUC: 0.82 Sensitivity: 0.98 Specificity: 0.53	Pretreatment morphological MRI radiomics signatures can predict early response to induction chemotherapy in patients with NPC
Zhang et al., 2017 (China)	Advanced NPC	Local and distant failure	MRI (CE-T1WI, T2WI)	110	70 training; 40 testing	MATLAB	RF, AdaBoost, & SVM	AUC: 0.85	Radiomics-based prediction of local failure and distant failure in advanced NPC, which could enhance the use of applications of radiomics in precision oncology and clinical practice
Zhang et al., 2017 (China)	Advanced NPC	Progression (LR or DM)	MRI (CE-T1WI, T2WI)	113	80 training; 33 testing	ITK-SNAP	LASSO	AUC: 0.89	MRI-based radiomics biomarkers present high accuracy in the pre-treatment prediction of progression in advanced NPC
Li et al., 2018 (China)	NPC	PFS	MRI (CE-T1WI, T2WI)	81	57 training; 24 testing	PyRadiomics	PCA, KNN, ANN, & SVM	Accuracies (ANN: 0.812, KNN: 0.775, SVM: 0.732).	Radiomics analysis can serve in finding imaging biomarkers to facilitate early salvage for NPC patients who are at risk of in-field recurrence
Lv et al., 2018 (China & USA)	NPC	Differential diagnosis	PET/CT	106	69 training; 37 testing	ITK-SNAP	LOOCV	–	Radiomics-generated features outperformed conventional metrics in differentiating NPC from chronic nasopharyngitis.
M.D. Anderson Cancer Center, Head and Neck Quantitative Imaging Working Group 2018 (USA)	OPSCC	Outcome prediction (local recurrence)	CT	465	225 training; 165 tuning; 45 testing	MATLAB (IBEX in MATLAB)	Bootstrap resampled recursive portioning analysis (RPA)	Discriminating value: 94 %	Radiomics features of the primary tumor can predict for local control in OPC patients
Ren et al., 2018 (China)	HNSCC	Tumor staging	CT	127	85 training; 42 testing	ITK-SNAP, MATLAB	LASSO & LR	Accuracy: 0.85	Radiomics signature based on MRI could discriminate stage I-II from stage III-IV HNSCC,

(*continued on next page*)

Table 2 (*continued*)

Studies									
									which may serve as a complementary tool for preoperative staging
Chen et a., 2019 (China & USA)	HNC	Outcome prediction (LNM)	PET/CT	59	41 training; 18 testing	–	SVM & CNN	Accuracy: 0.88	Deep learning radiomics provides a more accurate way of predicting LNM using PET and CT.
Du et al., 2019 (China & USA)	NPC	Differential diagnosis	PET/CT	76	70:30	ITK-SNAP	FSCR + KNN, FSCR + SVM, RBF + SVM, FSCR + RF, mRMR, & RBF + SVM	AUC: 0.867 – 0.892)	ML-based radiomics could enhance precision diagnosis of NPC
Forghani, 2019 (Canada)	Oral cavity	Outcome prediction (Cervical LNM)	CT	87	70 % training; 30 % testing	– (in-house built program on MATLAB	RF & DNN	Accuracy: 88.0 %	Dual-energy CT texture analysis with machine learning can enhance non-invasive diagnostic tumor evaluation
Liu et al., 2019 (China)	Stage I-IVB NPC	Acute xerostomia	CT	45	–	3D Slicer	Linear regression	MSE: 0.90 R^2 score: 0.74	Radiation-induced acute xerostomia level could be early predicted based on the saliva amount and radiomics changes of the parotid glands during intensity-modulated radiotherapy delivery
Liu et al., 2019 (China)	Thyroid carcinoma	LNM	US	450	2:1	MATLAB	SVM	AUC: 0.73 Accuracy: 0.71	The radiomics model showed a promising method for assessing the risk of papillary thyroid carcinoma metastasis noninvasively
Ming et al., 2019 (China)	NPC	OS, DFS, LRFS, DMFS	MRI (CE-T1WI)	303	200 training; 103 testing	MATLAB	LASSO & LOOCV	C-index: 0.85	The combination of clinical and radiomics features can provide more information for precise treatment decision
Wu et al., 2019 (China)	HNSCC	Differentiation of HNSCC	CT	206	–	MATLAB, Python	Kernel PCA & RF	AUC: 0.96 Accuracy: 0.92	The CT-based radiomics signature could discriminate between well-differentiated, moderately differentiated, and poorly differentiated HNSCC and might serve as a biomarker for preoperative grading
Yu et al., 2019 (Hong Kong & China)	Stage II-IVB NPC	Predicting adaptive radiation therapy eligibility before treatment	MRI (CE-T1WI, T2WI)	70	51 training; 19 testing	3D Slicer	LASSO & LR	AUC: 0.85	MRI-based radiomics features showed promising results for pre-treatment identification of adaptive radiotherapy eligibility in NPC patients
Zhang et al., 2019 (China)	NPC	Outcome prediction (Assessment of distant metastasis)	MRI	176	123 training; 53 testing	PyRadiomics	mRMR & LASSO	AUC: 0.79	Radiomics can serve as a visual prognostic tool for distant metastasis prediction in NPC, and it can improve treatment decisions by aiding in the differentiation of patients with high and low risks of distant metastasis.
Zhou et al., 2019 (China)	NPC	Risk score	MRI (CE-T1WI,	658	424 training; 234 testing	MATLAB	k-means	C-index: 0.84	The quantitative multi-modalities MRI image

(*continued on next page*)

Table 2 (*continued*)

Studies									
			T2WI, T1WI						phenotypes reveal distinct survival subtypes
Akram et al., 2020 (Singapore)	NPC	Recurrent and non-recurrent regions	MRI (T1WI	14	–	Open source	–	–	The radiomics features extracted from pre-treatment MRI can potentially reflect the difference between recurrent and non-recurrent regions within a tumor
Feng et al., 2020 (China)	NPC	Clinical staging	PET/MRI	100	70 training 30 testing	Artificial Intelligence Kit (AK)	mRMR, LASSO, & LR	In MRI: AUC: 0.85 Accuracy: 0.83 Sensitivity: 0.75 Specificity: 0.86 In PET: AUC: 0.84 Accuracy: 0.75 Sensitivity: 0.90 Specificity: 0.69	The PET and MRI radiomics models were helpful in the diagnosis of NPC staging
Guo et al., 2020 (China)	LHSCC	Determine whether CT radiomics could aid in the prediction of thyroid cartilage invasion from laryngeal and hypopharyngeal cancer	CT	265	–	Radcloud Platform & Anaconda3 platform	LASSO, LR, & SVM	The radiomics-based models (AUC 0.905, 95 % CI 0.863–0.937) were more accurate than a clinical radiologist alone (0.721, 95 %CI: 0.663–0.774) in predicting thyroid cartilage invasion	Models based on CT radiomics features can improve the accuracy of predicting thyroid cartilage invasion from LHSCC and provide a new potentially noninvasive method for preoperative prediction of thyroid cartilage invasion from LHSCC.
Haider et al., 2020 (USA, Germany & Canada)	HPV-associated OPSCC	Outcome prediction (PFS & OS)	FDG-PET/CT (Local + TCIA)	311	235 HPV-positive and 76 HPV-negative OPSCC for testing	3D Slicer, PyRadiomics	RF	For PFS: C-index: 0.62	Radiomics imaging features extracted from pre-treatment PET/CT may provide complimentary information to the current AJCC staging scheme for survival prognostication and risk-stratification of HPV-associated OPSCC
Haider et al., 2020 (USA, Germany & Canada)	OPSCC & MLNs	Outcome prediction (HPV status)	PET/CT (Local + TCIA)	435 OPSCC 741 MLNs	For OPSCC: (326 for training; 109 for validation) For MLNs: (518 for training; 223 for validation)	3D Slicer	RF	AUC: 0.78	PET-based radiomics signatures yielded similar classification performance to CT-based models, with potential added value from combining PET- and CT-based radiomics for prediction of HPV status
Ho et al., 2020 (Taiwan)	HNSCC	Classify benign and malignant tumors, differentiate ENE	MRI	130	–	3D Slicer, Segmentation Wizard	Adam optimization algorithm	84.0% for differentiation between malignant and non-malignant 77 % for differentiating between ENE and non-ENE	This computer aided diagnosis can assist in clinical decision-making
Li et al., 2020 (China)	Thyroid cancer	LNM	US	126	–	ITK-SNAP, Ultrosomics, SPSS	LASSO, PCA, KNN, LR,	AUC: 0.80 Sensitivity: 0.73	Radiomics analysis has promising value in screening meaningful ultrasound features

(*continued on next page*)

Table 2 (*continued*)

Studies									
							SVM, AdaBoost, Bagging, RF, Extreme RF, DT, NB, & Gradient Boosting DT	Specificity: 0.80	in thyroid cancer patients with LNM
Mukherjee et al., 2020 (USA)	HNSCC	Tumor grading, extracapsular spread, perineural invasion, & HPV status	CT	184	113 training; 71 testing	MATLAB	PCA, LR, LASSO, & Clustering	AUC: 0.75 (tumor grade) AUC:	Radiomics CT models have the potential to predict characteristics typically identified on pathologic assessment of HNSCC
Romeo et al., 2020 (Italy)	Oral cavity squamous cell carcinoma, oropharynx	Outcome prediction (Tumor grade and nodal status)	CT	40	75 % of the data for training; 25 % for testing	ITK-SNAP, 3D slicer	J48 algorithm	For tumor grade: Accuracy: 92.9 % For nodal status: 90.0 %	A radiomics ML approach applied to primary tumor lesions can predict tumor grade and nodal status in patients with OC and OP SCC
Tran et al., 2020 (Canada)	HNC	Outcome prediction (Measuring response to radiotherapy)	PET/CT and MRI	36	–	MATLAB	k-NN & NB	Accuracy > 80 %	Quantitative ultrasound-radiomics can predict radiotherapy response at 3 months as early as 24 h with reasonable accuracy
Zhang et al., 2020 (China)	NPC	Radiotherapy-induced temporal lobe injury	MRI (CE-T1WI, T2W1)	242	–	ITK-SNAP, MATLAB	RF	AUC: 0.83	The three developed radiomics models can dynamically predict RTLI in advance, enabling early detection and allowing clinicians to take preventive measures to stop or slow down the deterioration of RTL
Fatima et al., 2021 (Canada)	HNSCC	Outcome prediction (Recurrence)	US	51	–	MATLAB	SVM & KNN	AUC: 0.75 Accuracy: 80 %	Quantitative ultrasound Delta-radiomics can predict higher risk of recurrence with reasonable accuracy in HNSCC
Fh et al., 2021 (Hong Kong)	HNSCC	Outcome prediction (Survival prediction)	CT	188	Leave-one-out cross validation	PyRadiomics, 3D Slicer, & MATLAB	LOOCV & CNN	For PTV: Accuracy: 77.7 % AUC: 0.934 (Death) AUC: 0.932 (recurrence). For GTV: Accuracy: 74.3 % AUC: 0.947 (Death) AUC: 0.956 (cancer recurrence)	Using both GTV and PTV radiomics features in the DL-ANN model can aid in predicting HNSCC-related prognosis and cancer recurrence
Haider et al., 2021 (USA, Germany & Canada)	HPV-associated OPSCC	Outcome prediction (LRP)	PET/CT	190	–	3D Slicer	RF	C-index (IQR): 0.76	This approach can provide novel quantitative imaging biomarkers for risk stratification and prediction of post-radiotherapy LRP in HPV-associated OPSCC.
Kim et l., 2021 (Republic of Korea)	NPC	PFS	MRI	81	57 training; 24 testing	3D Slicer , PyRadiomics, R package	LASSO & Cox regression	AUC: 0.76 – 0.81	Integration of MR-based radiomics features with clinical and stage variables improved

(*continued on next page*)

Table 2 (*continued*)

Studies									
									the prediction PFS in patients diagnosed with NPC
Peng et al., 2021 (China)	Stage III-IVB NPC	LR & DM	PET/CT	85	–	ITK-SNAP	SVM, LOOCV, & Sequential floating forward selection	AUC value of 0.8290 (sensitivity: 0.8438, specifcity: 0.7736	Locoregional recurrence (LR) and DM of locally advanced NPC can be predicted using radiomics analysis of pretreatment [18F] FDG PET/CT
Wang et al., 2021 (China)	Tongue cancer	Outcome prediction (LNM)	MRI	236	157 training; 79 testing	ITK-SNAP, AIMT	PCA & SVM	AUC: 0.872	Radiomics analysis with a 10-mm peritumoral extension had excellent power to predict LNM and prognosis in tongue cancer
Woolen et al., 2021 (USA)	Laryngeal and hypopharyngeal cancer	Outcome prediction (DFS)	CT	44	–	Perfusion-4, ROCKIT	LDA & Two-loop leave-one-out	AUC: 0.69	CT perfusion and radiomics features in combination form potential predictors of one-year disease free survival in laryngeal and hypopharyngeal cancer patients.
Yang et al., 2021 (China)	NPC	Induction chemotherapy response	CT	297	208 training; 89 testing	PyRadiomics	ResNet50	AUC: 0.81 Accuracy: 68.5 %	The noninvasive deep learning method could provide efficient prediction of treatment response to induction chemotherapy in locally advanced NPC and might be a practicable approach in therapeutic strategy decision-making
Zhang et al., 2021 (USA)	LAHNC	Outcome prediction (lymph node response to induction therapy	CT	27	3:1	3D Slicer	LASSO Regression model	AUC: 0.75	A pretreatment CT-based lymph node radiomics signature combined with clinical parameters was able to predict nodal response to induction chemotherapy for patients with locally advanced HNSCC
Zhong et al., 2021 (United Kingdom)	Larynx/ Hypopharyngeal	Outcome prediction (early-disease progression)	PET/CT	72	–	LIFEx	RF	AUC: 0.94	FDG PET-CT derived radiomics features are potential predictors of early disease progression in patients with locally advanced larynx and hypopharynx SCC
42 Kaźmierska et al., 2022 (Poland & Canada)	LAHNC	Outcome prediction (Incomplete response and disease progression)	CT	290	–	PyRadiomics	LASSO LR	AUROC: 0.68	Combining clinical and radiomics features did not improve model's performance
Kim et al., 2022 (Republic of Korea)	HNSCC	Outcome prediction (local tumor recurrence)	MRI	215	161 training; 54 testing	MATLAB	LASSO LR	AUC: 0.77	The radiomics model using apparent diffusion coefficient maps exhibited higher diagnostic performance than those of the radiomics models using T2WI or CE-T1WI and clinical parameters in the diagnosis of local tumor recurrence in HNSCC following definitive treatment.
Nakajo et al., 2022 (Japan)	Hypopharyngeal Cancer	Outcome prediction (PFS)	FDG-PET/CT	100	–	LIFEx	LR NN	AUC: 0.80 Accuracy: 0.70	Radiomics-based features were useful for predicting prognosis

(*continued on next page*)

Table 2 (*continued*)

Studies									
							KNN NB SVM		in patients with hypopharyngeal cancer
Xi et al., 2022 (China)	NPC	Outcome prediction (response to induction chemoradiotherapy)	MRI	272	155 training; 66 testing; and 51 external validation	Artificial Intelligence Kit (AK)	LASSO LR	AUC: 0.86 Accuracy: 0.82	The pretreatment MRI radiomics model and pre- and post-IC Delta radiomics models could predict the IC-CCRT response of NPC in non-epidemic areas
Lin et al., 2023 (China)	NPC	Outcome prediction (Cervical LNM)	US	205 (NPC Cervical LNM) 284 Benign lymphadenopathy	7:3	mRMR & LASSO	LR	AUC: 0.88	Ultrasound radiomics analysis has potential value in screening meaningful ultrasound features and improving the diagnostic efficiency of ultrasound in cervical LNM of patients with NPC

ACM: All-cause mortality; **AJCC:** American Joint Committee on Cancer; **AUC:** Area Under Curve; **AUROC:** Area Under Receiving Operating Characteristics Curve; **c-index:** Concordance Index; **CN:** Convolution Network; **CNN:** Convolution Neural Network; **CT:** Computed Tomography; **DFS:** Disease-free survival; **DL-ANN:** Deep learning Artificial Neural Network; **DM:** Distant Metastasis; **DMFS:** Distant Metastasis Free-Survival; **DNN:** Deep Neural Network; **DT:** Decision Tree; **EBV:** Epstein Barr Virus; **ENE:** Extra Nodal Extension; **FDG:** Fluoro-2-deoxy-D-glucose; **GTV:** Gross Tumor Volume; **HNC:** Head and Neck cancer; **HNSCC:** Head and Neck Squamous Cell Carcinoma; **HPV:** Human papillomavirus; **KNN**: K-Nearest Neighbors; **LASSO**: Least Absolute Shrinkage and Selection Operator; **LDA**: Linear Discriminant Analysis; **LF**: Local Failure; **LR**: Logistic Regression; **LAHNC:** Locally Advanced Head and Neck Cancer; **LHSCC:** Laryngeal and hypopharyngeal squamous cell carcinoma; **LRP**: Locoregional Disease Progression; **LRFS**: Locoregional Recurrence Free Survival; **LNM:** Lymph Node Metastasis; **LOOCV:** Leave One Out Cross Validation; **IQR:** Interquartile Range; **mRMR:** Maximum relevance minimum redundancy; **ML:** Machine learning; **MLN:** Metastatic Lymph Nodes; **MRI:** Magnetic Resonance Imaging; **MSE:** Mean Squared Error; **NB:** Naïve Bayes; **NPC:** Nasopharyngeal Cancer; **OPC:** Oropharyngeal Cancer; **OPSCC**: Oropharyngeal Squamous Cell Carcinoma; **OS**: Overall Survival; **PCA:** Principal Component Analysis; **PET:** Positron Emission Tomography; **PFS**: Progression Free Survival; **PTV**: Planning Target Volume; **RBF:** Radial Basis Function; **RF:** Random Forest; **SVM**: Support Vector Machine; **USA:** United States of America.

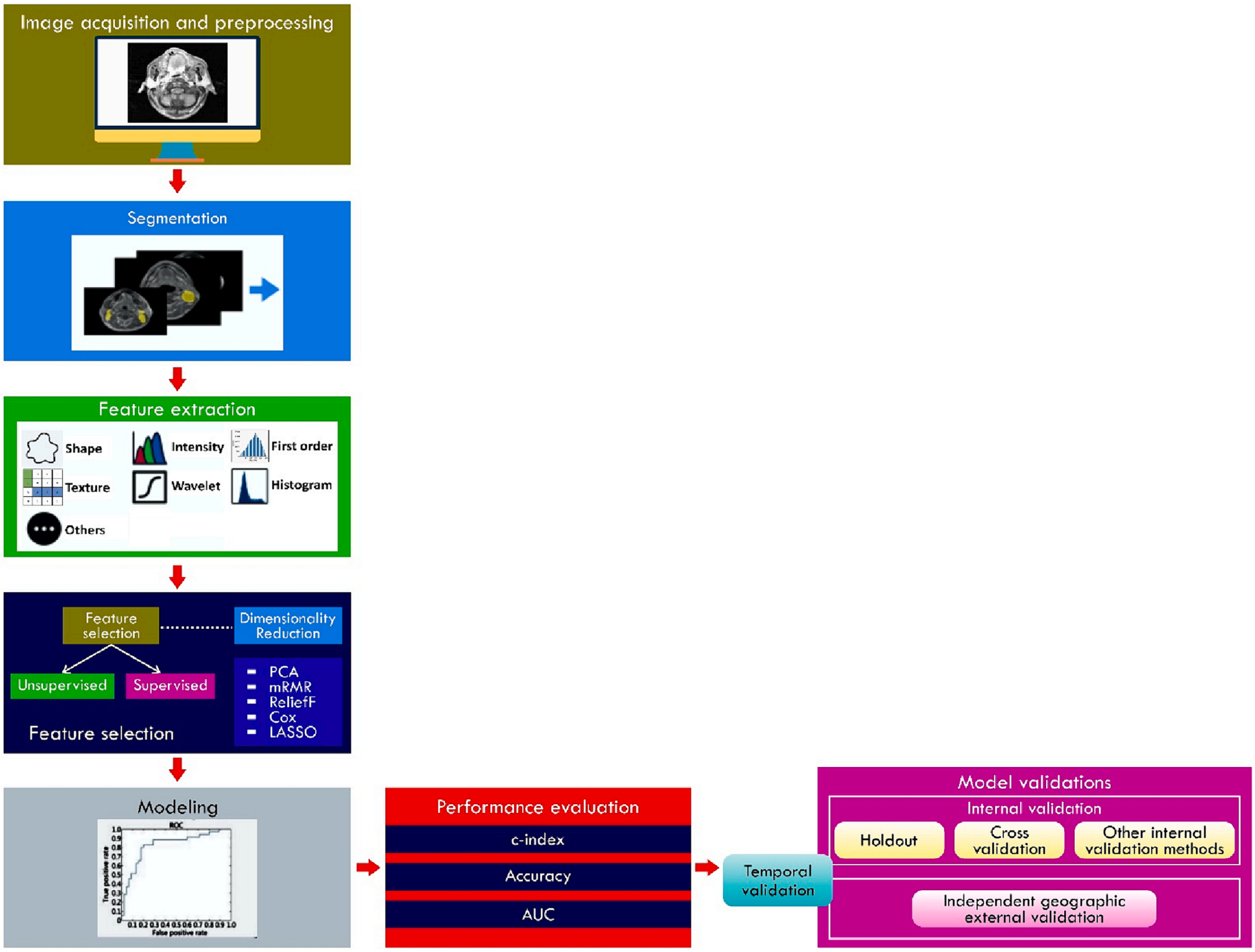

Fig. 2. A typical workflow of an artificial intelligence-based radiomics.

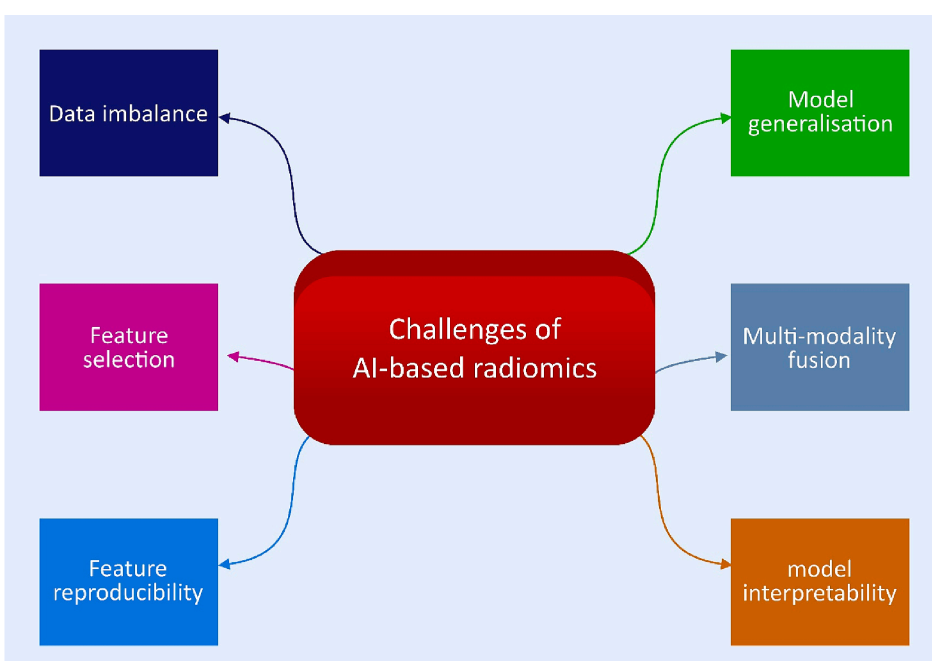

Fig. 3. Some challenges of artificial intelligence-based radiomics.

Table 3
Tabular presentation of PROBAST results.

Study	ROB				Applicability			Overall	
	Participants	**Predictors**	**Outcome**	**Analysis**	**Participants**	**Predictors**	**Outcome**	**ROB**	**Applicability**
Folkert et al., 2017 (USA)	+	+	+	+	+	+	+	+	+
Vallieres et al., 2017 (Canada)	+	+	+	+	+	+	+	+	+
Wang et al., 2017 (China & USA)	+	+	?	+	+	+	+	?	+
Zhang et al., 2017 (China)	+	+	+	+	+	+	+	+	+
Zhang et al., 2017 (China)	+	+	+	+	+	+	+	+	+
Li et al., 2018 (China)	+	+	+	+	+	+	+	+	+
Lv et al., 2018 (China & USA)	+	+	−	+	+	+	+	−	+
M.D. Anderson Cancer Center, Head and Neck Quantitative Imaging Working Group 2018 (USA)	+	+	+	+	+	+	+	+	+
Ren et al., 2018 (China)	+	+	+	+	+	+	+	+	+
Chen et a., 2019 (China & USA)	+	+	+	+	+	+	+	+	+
Du et al., 2019 (China & USA)	+	+	+	+	+	+	+	+	+
Forghani, 2019 (Canada)	+	+	+	+	+	+	+	+	+
Liu et al., 2019 (China)	+	+	?	+	+	+	+	?	+
Liu et al., 2019 (China)	+	+	+	+	+	+	+	+	+
Ming et al., 2019 (China)	+	+	+	+	+	+	+	+	+
Wu et al., 2019 (China)	+	+	?	+	+	+	+	?	+
Yu et al., 2019 (Hong Kong & China)	+	+	+	+	+	+	+	+	+
Zhang et al., 2019 (China)	+	+	+	+	+	+	+	+	+
Zhou et al., 2019 (China)	+	+	+	+	+	+	+	+	+
Akram et al., 2020 (Singapore)	+	+	?	?	+	?	?	?	?
Feng et al., 2020 (China)	+	?	+	+	+	?	+	?	?
Guo et al., 2020 (China)	+	?	+	+	+	+	+	?	+
Haider et al., 2020 (USA, Germany & Canada)	+	+	+	+	+	+	+	+	+
Haider et al., 2020 (USA, Germany & Canada)	+	+	+	+	+	+	+	+	+
Ho et al., 2020 (Taiwan)	+	?	+	+	+	+	+	?	+
Li et al., 2020 (China)	+	?	+	+	+	+	+	?	+
Mukherjee et al., 2020 (USA)	+	+	+	+	+	+	+	+	+
Romeo et al., 2020 (Italy)	+	+	+	+	+	+	+	+	+
Tran et al., 2020 (Canada)	+	?	+	+	+	+	+	?	+
Zhang et al., 2020 (China)	+	?	+	+	+	+	+	?	+
Fatima et al., 2021 (Canada)	+	?	+	+	+	+	+	?	+
Fh et al., 2021 (Hong Kong)	+	+	+	+	+	+	+	+	+
Haider et al., 2021 (USA, Germany & Canada)	+	?	+	+	+	+	+	?	+
Kim et l., 2021 (Republic of Korea)	+	+	+	+	+	+	+	+	+
Peng et al., 2021 (China)	+	?	+	+	+	+	+	?	+
Wang et al., 2021 (China)	+	+	+	+	+	+	+	+	+
Woolen et al., 2021 (USA)	+	?	+	+	+	+	+	?	+
Yang et al., 2021 (China)	+	+	+	+	+	+	+	+	+
Zhang et al., 2021 (USA)	+	+	+	+	+	+	+	+	+
Zhong et al., 2021 (United Kingdom)	+	?	+	+	+	+	+	?	+
Kaźmierska et al., 2022 (Poland & Canada)	+	?	+	+	+	+	+	?	+
Kim et al., 2022 (Republic of Korea)	+	+	+	+	+	+	+	+	+
Nakajo et al., 2022 (Japan)	+	?	+	+	+	+	+	?	+
Xi et al., 2022 (China)	+	+	+	+	+	+	+	+	+
Lin et al., 2023 (China)	+	+	+	+	+	+	+	+	+

PROBAST = Prediction model Risk Of Bias Assessment Tool; ROB = Risk of Bias.
+ Indicates Low ROB/Low concern regarding applicability.
− Indicates High ROB/high concern regarding applicability.
? Indicates unclear ROB/unclear concern regarding applicability.

3. Results

3.1. Results of the database search

A total of 659 hits were retrieved. After deleting duplicates (N = 168), and irrelevant papers (N = 446), we found 45 studies eligible to be included in this review as shown in Fig. 1 [7,24–67].

3.2. Characteristics of relevant studies

All the 45 articles included in this review were published in English. Twenty-five of them were conducted in Asia [7,27,28,32,36–45,48,49,53,56–58,60,64–67], 9 were conducted in America [24,25,31,35,50,52,54,59,61], 2 in Europe [51,62], and 9 were multicenter studies [26,30,33,34,40,46,47,55,63] (Table 2). All but two of the included studies showed high-quality applicability and low risk of bias (Tables 3-4) [43,44]. Similarly, all the included studies showed a quality that ranged from 84 – 87 % (Supplementary I for TRIPOD assessment). Five studies were conducted in the year 2017 [24–28], 4 studies in 2018 [29–32], 10 in 2019 [33–42], 11 studies in 2020 [43–53], 10 studies in 2021 [7,54–62], and the remaining 5 studies in the year 2022 [63–67].

3.3. Current application of AI-based radiomics in HNC oncology

The findings of the published studies (Table 2) suggest that the application of AI-based radiomics in head and neck oncology includes – radiomics-based cancer diagnostics [30,32,34,39,44,45,48,50,51] and radiomics-based cancer prognosis [7,24–29,31,33,35–38,40–43,46,47,49,52–67] (Fig. 4). The radiomics-based cancer diagnostics includes tumor staging [32,44,51], tumor grading [39,50], HPV status [47,50], and classification of malignant and benign tumors [30,34,39,45,48]. Similarly, radiomics-based cancer prognosis covers predictions of

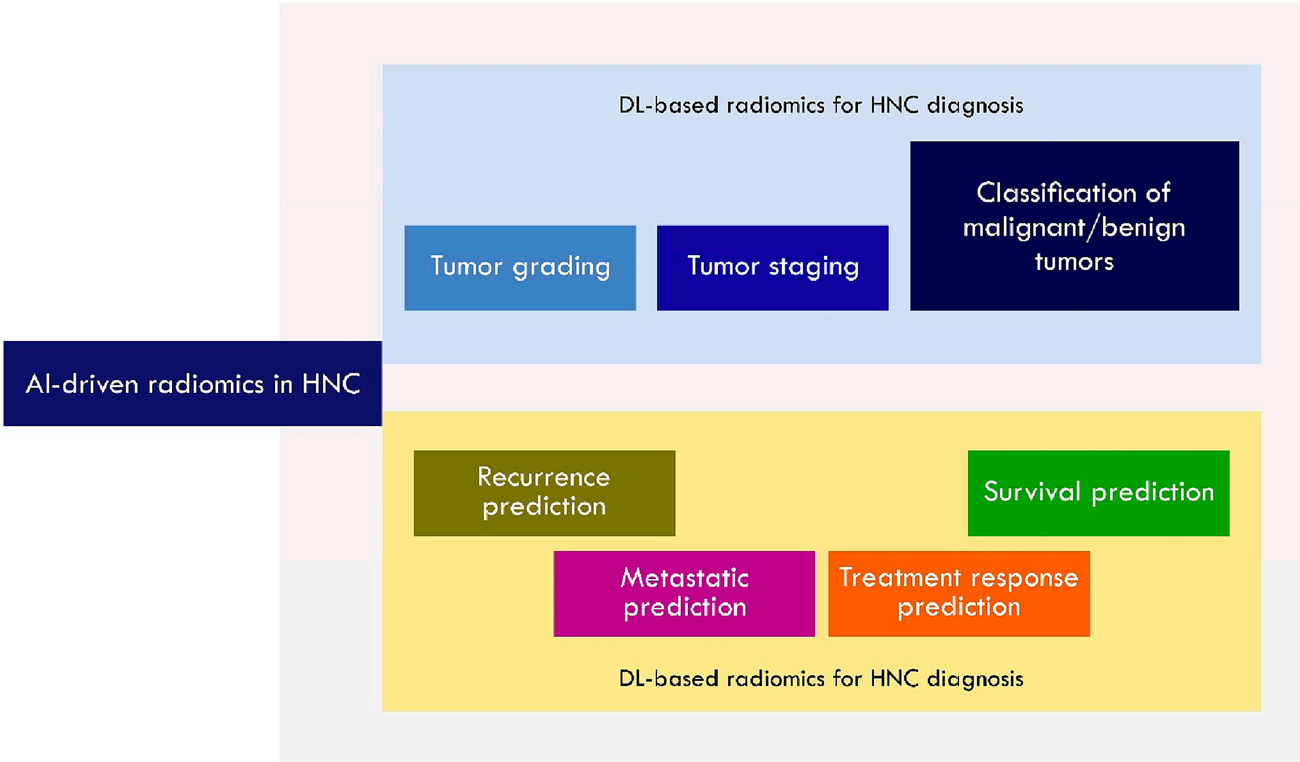

Fig. 4. Applications of artificial intelligence-based radiomics.

metastasis [33,35,37,41,49,51,57,58,62,63,67], local and distant failures [27], treatment response [26,28,36,38,40,52,53,60,61,66], recurrence [29,31,54,55,57,64], and survival [7,24,25,42,46,56,59,65] (Fig. 4).

3.4. Workflow of AI-based radiomics in HNC oncology

The typical workflow highlighted in this review includes image acquisition and preprocessing, segmentation, feature engineering (extraction and selection), machine or deep learning modeling, performance evaluation, and model validations (Fig. 2).

3.5. Software used for AI-based radiomics in HNC oncology

Some of the commercial/open-source software used for feature extraction in radiomics include 3D Slicer, MATLAB, PyRadiomics, ITK-SNAP, Python, R, Artificial Intelligence Kit, Perusion-4, and ROCKIT (Fig. 5). Similarly, various AI (ML or DL) models have been proposed for feature selection. Some of these include Least Absolute Shrinkage and Selection Operator (LASSO), Leave One Out Cross Validation (LOOCV), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Forest (DF), and Convolutional Neural Network (CNN) (Fig. 5).

3.6. Challenges, limitations, and concerns of AI-based radiomics in HNC oncology

This review reveals some limitations affecting the recommendation of AI-based models for further clinical evaluations. These include data imbalance, feature selection, feature reproducibility, model generalization, multi-modality fusion and model interpretability (Fig. 3).

4. Discussion

This review highlights the applications of artificial intelligence-based radiomics (AI-based radiomics) in head and neck cancer (HNC). Furthermore, it highlights the workflow of AI-based radiomics, challenges, and limitations impeding the adoption of AI-based radiomics in clinical practice. Finally, it suggests possible solutions and areas of future research to ensure that the potential of AI-based radiomic model as an auxiliary tool in HNC management is actualized in daily clinical practice.

4.1. Current applications of AI-based radiomics for HNC management

4.1.1. ML and deep learning-based radiomics for HNC diagnosis

The traditional diagnosis and treatment approach for HNC patients has seemingly improved over the years. However, there are possibilities of misdiagnoses and missed diagnoses that may hamper early intervention aimed at effective management of HNC. Thus, decreasing the patient's survival rates. One of the approaches to enhance the effective management of HNC is to leverage the potential of AI. In recent years, specifically in terms of the intervention of AI as an ancillary tool for the management of HNC, radiomics has been reported to assist clinicians in insightful tumor staging [32,44,51], tumor grading [39,50], and classification of malignant and benign tumors [30,34,39,45,48]. DL-based radiomics uses feature engineering to detect intra-tumoral properties that are usually undetected during visual inspection in clinical images [68].

DL-based radiomics for tumor grading offers an approach for describing the magnitude of head and neck tumor atypia [39,50,69]. Several studies have emphasized the potential of PET/CT-based radiomics features to develop a DL-based radiomics model for HNC tumor grading [39,50]. This grading of tumors based on DL-based radiomics models can provide vital information on the patients that will allow the

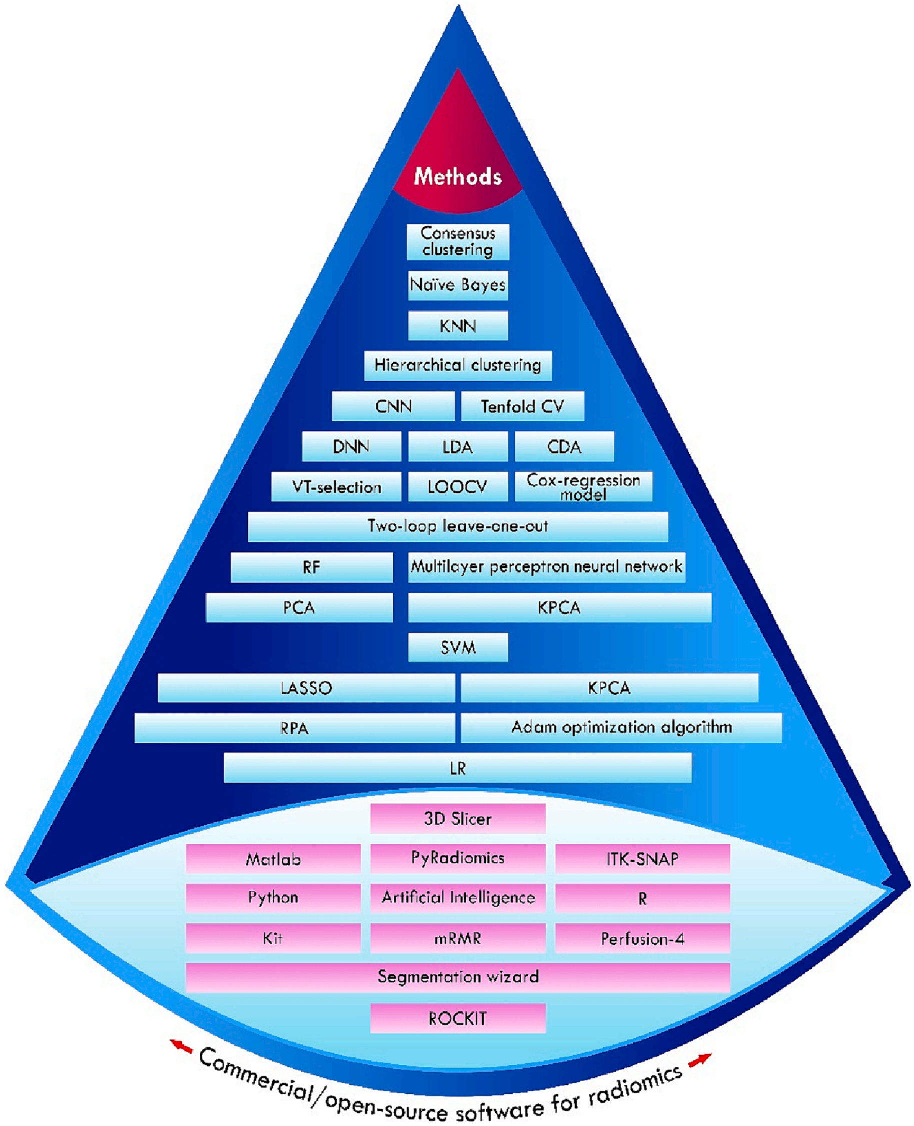

Fig. 5. Some open-software and commercial applications for feature engineering and radiomics model development. (Artificial Intelligence for Artificial Intelligence Kit and Kit for open source applications with radiomics ability)

clinicians to map out an effective treatment plan for the patients to reduce the possibility of recurrence and treatment toxicities.

Besides knowing the magnitude of head and neck tumor atypia, the information about the severity of HNC in terms of its size and degree of its spread within HNC patients' bodies is also vital for the management of this type of cancer. In clinical practice, histologic grading of cancer is typically done after imaging and pathological biopsies. DL-based radiomics model has the potential to aid in preoperative staging of the tumor [32]. For example, the study by Ren et al. used radiomics for preoperative staging of HNC based on MRI [32]. Their study found that radiomics signature based on MRI could discriminate stage I-II from stage III-IV HNSCC, which may serve as a complementary tool for pre-operative staging [32].

In the era of personalized and precision oncology, early diagnosis becomes imperative. AI approaches such as DL-based radiomics is poised to serve as an auxiliary decision tool for clinicians. For instance, the study by Ho et al. identified some features (n = 89) within MRI images that can be used to differentiate between benign and malignant lymph nodes [48]. In the quest for early intervention in HNC management, this can serve as a diagnostic tool.

4.1.2. ML and Deep learning-based radiomics for HNC prognostication

Numerous studies have highlighted promising results regarding the application of DL-based radiomics for various outcome prognosis in HNC. Fh et al., used planning target volumes and gross tumor volumes to predict survival and tumor recurrence in HNC [7]. Similarly, for oropharyngeal squamous cell carcinoma (OPSCC), which is a subsite within the head and neck region, Haider et al., explored the potential added value of baseline positron emission tomography (PET)/computed tomography (CT) radiomics features for prognosis and risk stratification of OPSCC [46]. They found out that radiomics-based model offers a promising approach for risk-stratification of HPV-associated OPSCC for progression free survival and overall survival [46]. The study concluded that radiomics imaging features extracted from pre-treatment PET/CT may provide complementary information to the current AJCC staging scheme for survival prognosis and risk-stratification of HPV-associated OPSCC [46]. Several other studies have also highlighted the significance of DL-radiomics for predict various forms of survivals – long-term, recurrence-free, disease-free, and overall survival [58,59,70,71]. Besides, there are other studies that specifically highlight DL-radiomics for the estimation of lymph node metastasis based on routine medical images (PET/CT) [35,41,58,67,72] and recurrence prediction [7,24,31,64].

Beyond the spectrum of treatment planning is the potential response of HNC patients to radical treatments – radiotherapy, chemotherapy, or chemoradiotherapy. Therefore, delta-radiomics, which is an extension to the field of radiomics examines the response of HNC patients to treatment sequelae. In recent years, studies are emerging on the use of DL-based radiomics for the prediction of treatment response after a specific period [52,61,63,66]. Considering the significant physical and psychological consequences of treatment side effects on HNC patients, follow-up studies are emerging on the prediction of adverse side effects of cancer treatments o tissues [73,74]. The significance of these types of studies are aimed at treatment optimization and personalized treatment plans.

4.2. Workflow for deep learning-based radiomics in HNC

Images form the cornerstone of radiomics studies. Therefore, the workflow for deep learning radiomics begins with image acquisition and preprocessing (Fig. 2). As shown in Table 2, these images include computed tomography (CT), positron emission tomography (PET), a combination of both CT and PET (PET/CT), magnetic resonance imaging (MRI), and ultrasound (US) imaging. Depending on the objectives of the study, CT imaging modality is mostly preferred because it is a less time-consuming, widely available, and less costly option, and most importantly, it offers repeatability and robustness of radiomics features [75]. PET, on the other hand, has been found to require significant harmonization efforts because it produces radiomics features that are susceptible to differences in reconstruction parameters [75,76]. However, MRI modality has been reported to be superior to CT in distinguishing exquisite soft tissue changes from cartilage abnormalities [68]. Additionally, it offers the ability to perform functional imaging at a high resolution [75]. Ultrasound imaging represents the least used imaging modality despite its less expensive nature because of its high inter-operator variability that may pose a barrier to reliable and reproducible radiomics applications [68,75,76].

The next step in the workflow, following pre-processing, is tumor contouring or annotation in 2D or 3D, that is, segmentation. The delineation of segmentation done in 2D is known as region of interest (ROI) while in 3D, it is called volume of interest (VOI). Although, in some cases, ROI may also be used interchangeably for 3D segmentation. Segmentation is usually done manually by clinicians or trained personnel such as researchers. This can be time-consuming and highly subjective to inter-operator variability because it depends significantly on the expert level of the personnel. Therefore, semi-automatic and automatic segmentation approaches such as deep learning via U-Net, adaptive Bayesian segmentation algorithms, and fuzzy c-means are gaining momentum in recent years.

Feature extraction, which is a purely software-based process follows the segmentation step. For simplicity, Gillies et al. classified features extracted from clinical images into two distinct groups – semantic and agnostic features [77]. Semantic features follow the computer-aided quantification of characteristics or terms that have biological justification (prognostic value) in radiology to describe ROI such as size, shape (sphericity, volume, and surface area), first-order (properties of histogram – entropy and mean and median value), location, presence of necrosis, etc. Nonetheless, the selection of ROI depends on the nature of the study [68]. Despite the high prognostic value attached to semantic features, the downsize includes higher inter-observer variability, slower throughput, and higher variance [75]. Agnostic features, on the other hand, have no biological justification because they entirely arose from the field of computer vision [75]. Hence, they are purely empirical and subjected to the prospect of false discovery due to statistical fluctuations [75]. Despite this, they offer a faster and larger (usually in hundreds or even thousands) number of extracted features.

In general, the quality of the images has an impact on the image features extracted. Thus, highlighting the image pre-processing step is important. The approaches of image pre-processing include image smoothing (averaging), image noise reduction (applying Gaussian filters), and image enhancements (histogram equalization, deblurring, and resampling) [75]. Of note, texture feature calculations would require interpolation (down-sampling or up-sampling) to isotropic voxel spacing to be rotationally invariant and for reproducibility. This would allow for comparison between different datasets. In addition, to make calculations of texture features tractable, discretization (fixed number of bins or bins with fixed bin width) of image intensities inside the ROI is often required [75].

Traditionally, hundreds of radiomics features are extracted from the feature extraction process. Using all these features for modeling would lead to model complexity, multicollinearity, and low generalizability. Hence, feature selection is important to select relevant features that have a maximum correlation with the endpoints and remove the redundant features. In general, feature selection involves filtering, embedded, and wrapper methods [68,78].

The actual predictive or prognostic model development begins after relevant features are selected. Once the quantitative phenotypic characteristics and other clinically available information about the tumor are computed from radiological images, statistical methods are usually used to construct a prediction model. However, in recent years, due to the advancement in technology and modifications to convolutional neural networks, classic ML or DL approaches have gained momentum for image analysis and construction of prediction models. The performance of AI-based radiomics models is evaluated using appropriate performance metrics. Usually, AI-based prognostic models can be accessed using Kaplan-Meier analysis, log-rank test, or concordance index (c-index) while AI-based radiomics models are evaluated using Receiver Operating Characteristic (ROC) curve metrics.

4.3. Concerns and limitations of deep learning-based radiomics in HNC

The studies presented in Table 2 showed promising results. However, there are certain limitations impeding the deployment of these models for further clinical assessment. These include data imbalance, feature selection, feature reproducibility, model generalizability, multi-modality fusion, and model interpretability (Fig. 3).

4.3.1. Data imbalance

One of the challenges affecting the generalizability of the trained model is class imbalance. Most datasets used for models usually have inherent imbalances in relation to the target endpoint. This leads to a biased model. Thus, resampling techniques (under-sampling or over-sampling) have been suggested to handle this concern [68]. In over-sampling, minority class data are copied to add more data while in under-sampling, large amounts of the majority class data are discarded. In either method, representation of the targeted endpoint is ensured for a viable model [68].

4.3.2. Feature selection

Considering hundreds or even thousands of radiomics features extracted from medical images, model complexity and overfitting of these features are a concern. However, there is no absolute rule-of-thumb to determine the minimum number of radiomics features required for training a model. Although, cross-validation has been touted to address this concern [13]. Nonetheless, in studies where the number of training samples is relatively small, cross-validation may not necessarily address the possibility of overfitting. Beyond the large number of radiomics features, most of these software-extracted features (agnostic features) are empirical with no biological justification. Therefore, it becomes a clinical concern how to explain the prognostic values of this type of extracted radiomics features.

4.3.3. Feature reproducibility

One of the major sources of concern relating to the widespread recommendation of AI-based models for further clinical assessment is the lack of feature replicability. While image acquisition and reconstructions step are well highlighted in the AI-based radiomics workflow, there are variations in terms of scan acquisition parameters, functional condition of the scanner (wear and tear and manufacturers and model variations), varying degrees of image enhancement on a particular scan, and variation in reconstruction parameters [68]. These factors have affected feature reproducibility to a significant extent. Similarly, manual segmentation by trained specialists such as radiologists or radiation-oncologists has introduced certain levels of interobserver variability due to varying experience and training [75].

4.3.4. Model generalizability

Most of the published AI-based radiomics studies were based on retrospective data. As a result, there may be an overestimation of the results in these studies because it remains unclear how these models would perform on unknown data or even prospective data [75]. Therefore, it is important to further validate these models for generalizability [15,79,80]. AI-based radiomics models can further be subjected to temporal and independent external validations. This will ensure the proper assessment of the behavior of the model on an independent dataset. In addition, the models should be evaluated on prospective data with the intention of conducting long-term follow-ups. This approach will further evaluate the true performance of the model for generalizability.

4.3.5. Multi-modality fusion

Considering the inherent vital information contained in different medical data formats, adding these formats to produce AI-based radiomics model may improve their diagnostic and prognostic abilities. To this end, different multi-modality fusions have emerged – early, intermediate, and late multimodality fusions. Early fusion is the fusion of multiple modalities such as radiomic CT, PET, MRI features, gene expressions, tumor-node-metastasis, clinical, clinic-radiologic, clinicopathologic, and biological features before the implementation of feature input classifier [81–85]. In intermediate fusion (inter-layer fusion), modalities between the input and output layers are fused during modeling [81,86]. In contrast to both early and intermediate fusions, late fusion, also known as decision-level fusion includes building a model by performing a certain processing fusion of different modalities to improve the model's performance [33]. However, despite these potential benefits of multi-modality fusion, assembling data from multiple centers poses potential legal and ethical regulations due to privacy and data ownership concerns. Distributed, federated, and transfer learning approaches are examples of emerging areas of research to address this concern [87].

4.3.6. Model interpretability

The issues regarding the interpretability of AI-based radiomics models remain a concern [79]. The concerns relating to software-enhanced radiomics features (agnostic features), and "black box" nature of classic ML and DL add to the issue of model interpretability. Although, for classic machine learning models, several efforts continue to emerge to produce local and global interpretations of the predictions made by the model [88]. However, for AI-based radiomics models, there are few emerging efforts to increase the interpretability problems of AI-based radiomics models.

4.4. Future prospects and study conclusion

Despite the challenges of multi-modality fusion, it remains one of the viable options for improved model performance aimed at personalized and precision oncology. Therefore, future studies on AI-based radiomics should explore any of the multi-modal fusion modalities to enhance the quality of the resulting model. This may include a combination of multiple imaging modalities, radiomics features with other types of features, or radio-genomics (radiomics with other omics approaches). For model generalizability, checklists, and guidelines for AI-based radiomics should be developed to guide the construction of these models.

In addition, the developed models should be properly validated with independent external validation as the gold standard performance of these models. Following independent external validations, these models should be evaluated on prospective cohorts and duly followed up prior to their recommendation for further clinical evaluation. Furthermore, the interpretability of these models can be considered at both feature and model levels, respectively. That is, the association between features and tumor heterogeneity can be explored to increase interpretability. Similarly, at the model level, local and global interpretations of the predictions made by these models should be explored to increase their interpretabilities. Future studies are warranted to compare the performance of radiomics generated features trained by statistical methods with those trained using machine or deep learning approach.

5. Study limitations

Admittedly, methodological limitations influenced our present review. Firstly, the included studies used varying medical image formats and model performance metrics. Therefore, we could not present a summarized performance of these models for each of the highlighted areas of application of AI-based radiomics. Secondly, not all the included studies assessed the performance of these models on an independent external validation dataset. Therefore, we could not analyze the generalizability of the reported models in these studies. In addition, our review has a time frame limit since we used a systematic literature search methodology. This means that the present analysis may miss any important studies that were reported after this time frame.

5.1. Conclusions.

In this systematic review, we examined AI-based radiomics (ML or DL) in HNC. In the first perspective, we focused on the application of these models for diagnostic and prognostic purposes. In the second perspective, we provided an overview of the AI-based radiomics workflow, as well as barriers affecting the widespread implementation of these models. Our review suggests that AI-based radiomics has undoubtedly great potential for improving diagnostic evaluation and care of HNC patients. By enabling the extraction of higher-level data that are currently largely under-utilized in routine clinical practice, AI-based radiomics has the potential to contribute to personalized and precision oncology. The multiple challenges presented in this review provide exciting areas for future research and development.

CRediT authorship contribution statement

Rasheed Omobolaji Alabi: Conceptualization, Data curation, Formal analysis, Methodology, Resources, Writing – original draft. **Mohammed Elmusrati:** Methodology, Supervision, Writing – review & editing. **Ilmo Leivo:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Alhadi Almangush:** . **Antti A. Mäkitie:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

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.

Acknowledgment

The Sigrid Jusélius Foundation (AM 230129). The Helsinki University Hospital Research Fund (TYH2024203). The Turku University Hospital Research Fund (11290).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2024.105464.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, Global cancer statistics,, et al., GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 2021 (2020), https://doi.org/10.3322/caac.21660.
[2] D.E. Johnson, B. Burtness, C.R. Leemans, V.W.Y. Lui, J.E. Bauman, J.R. Grandis, Head and neck squamous cell carcinoma, Nat. Rev. Dis. Primers 6 (2020) 92, https://doi.org/10.1038/s41572-020-00224-3.
[3] M. Gormley, G. Creaney, A. Schache, K. Ingarfield, D.I. Conway, Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors, Br. Dent. J. 233 (2022) 780–786, https://doi.org/10.1038/s41415-022-5166-x.
[4] E.M. Sturgis, P.M. Cinciripini, Trends in head and neck cancer incidence in relation to smoking prevalence: An emerging epidemic of human papillomavirus-associated cancers? Cancer 110 (2007) 1429–1435, https://doi.org/10.1002/cncr.22963.
[5] M.C. Yu, J.-M. Yuan, Epidemiology of nasopharyngeal carcinoma, Semin. Cancer Biol. 12 (2002) 421–429, https://doi.org/10.1016/S1044579X02000858.
[6] M. Tortora, L. Gemini, A. Scaravilli, L. Ugga, A. Ponsiglione, A. Stanzione, et al., Radiomics Applications in Head and Neck Tumor Imaging: A Narrative Review, Cancers 15 (2023) 1174, https://doi.org/10.3390/cancers15041174.
[7] Fh T, Cyw C, Eyw C. Radiomics AI prediction for head and neck squamous cell carcinoma (HNSCC) prognosis and recurrence with target volume approach. BJR|Open 2021;3:20200073. https://doi.org/10.1259/bjro.20200073.
[8] H.J.W.L. Aerts, E.R. Velazquez, R.T.H. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, et al., Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, Nat. Commun. 5 (2014) 4006, https://doi.org/10.1038/ncomms5006.
[9] Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. JAMA Oncol 2017;3:524. https://doi.org/10.1001/jamaoncol.2016.5688.
[10] R.O. Alabi, A. Almangush, M. Elmusrati, A.A. Mäkitie, Deep Machine Learning for Oral Cancer: From Precise Diagnosis to Precision Medicine, Front Oral Health 2 (2022) 794248, https://doi.org/10.3389/froh.2021.794248.
[11] Alabi RO, Bello IO, Youssef O, Elmusrati M, Mäkitie AA, Almangush A. Utilizing Deep Machine Learning for Prognostication of Oral Squamous Cell Carcinoma—A Systematic Review. Frontiers in Oral Health 2021;2. https://doi.org/10.3389/froh.2021.686863.
[12] R.O. Alabi, M. Elmusrati, I. Sawazaki-Calone, L.P. Kowalski, C. Haglund, R. D. Coletta, et al., Machine learning application for prediction of locoregional recurrences in early oral tongue cancer: a Web-based prognostic tool, Virchows Arch. 475 (2019) 489–497, https://doi.org/10.1007/s00428-019-02642-5.
[13] R.O. Alabi, M. Elmusrati, I. Sawazaki-Calone, L.P. Kowalski, C. Haglund, R. D. Coletta, et al., Comparison of supervised machine learning classification techniques in prediction of locoregional recurrences in early oral tongue cancer, Int. J. Med. Inf. 104068 (2019), https://doi.org/10.1016/j.ijmedinf.2019.104068.
[14] R.O. Alabi, A.A. Mäkitie, M. Pirinen, M. Elmusrati, I. Leivo, A. Almangush, Comparison of nomogram with machine learning techniques for prediction of overall survival in patients with tongue cancer, Int. J. Med. Inf. 145 (2021) 104313, https://doi.org/10.1016/j.ijmedinf.2020.104313.
[15] R. Alabi, A. Almangush, M. Elmusrati, I. Leivo, A.A. Mäkitie, An interpretable machine learning prognostic system for risk stratification in oropharyngeal cancer, Int. J. Med. Inf. 168 (2022) 104896, https://doi.org/10.1016/j.ijmedinf.2022.104896.
[16] M. Tortora, L. Gemini, I. D'Iglio, L. Ugga, G. Spadarella, R. Cuocolo, Spectral Photon-Counting Computed Tomography: A Review on Technical Principles and Clinical Applications, J Imaging 8 (2022) 112, https://doi.org/10.3390/jimaging8040112.
[17] W. Duan, B. Xiong, T. Tian, X. Zou, Z. He, L. Zhang, Radiomics in Nasopharyngeal Carcinoma, Clin Med Insights Oncol 16 (2022) 117955492210791, https://doi.org/10.1177/11795549221079186.
[18] Q. Xie, X. Wang, J. Pei, Y. Wu, Q. Guo, Y. Su, et al., Machine Learning-Based Prediction Models for Delirium: A Systematic Review and Meta-Analysis, J. Am. Med. Dir. Assoc. 23 (2022) 1655–1668.e6, https://doi.org/10.1016/j.jamda.2022.06.020.
[19] G.S. Collins, J.B. Reitsma, D.G. Altman, K.G.M. Moons, Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement, BMJ 350 (2015) g7594–g, https://doi.org/10.1136/bmj.g7594.
[20] K.G.M. Moons, J.A.H. de Groot, W. Bouwmeester, Y. Vergouwe, S. Mallett, D. G. Altman, et al., Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist, PLoS Med. 11 (2014) e1001744.
[21] G.S. Collins, P. Dhiman, C.L. Andaur Navarro, J. Ma, L. Hooft, J.B. Reitsma, et al., Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence, BMJ Open 11 (2021) e048008.
[22] R.F. Wolff, K.G.M. Moons, R.D. Riley, P.F. Whiting, M. Westwood, G.S. Collins, et al., PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies, Ann. Intern. Med. 170 (2019) 51, https://doi.org/10.7326/M18-1376.
[23] K.G.M. Moons, R.F. Wolff, R.D. Riley, P.F. Whiting, M. Westwood, G.S. Collins, et al., PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration, Ann. Intern. Med. 170 (2019) W1, https://doi.org/10.7326/M18-1377.
[24] M.R. Folkert, J. Setton, A.P. Apte, M. Grkovski, R.J. Young, H. Schöder, et al., Predictive modeling of outcomes following definitive chemoradiotherapy for oropharyngeal cancer based on FDG-PET image characteristics, Phys. Med. Biol. 62 (2017) 5327–5343, https://doi.org/10.1088/1361-6560/aa73cc.
[25] M. Vallières, E. Kay-Rivest, L.J. Perrin, X. Liem, C. Furstoss, H.J.W.L. Aerts, et al., Radiomics strategies for risk assessment of tumour failure in head-and-neck cancer, Sci. Rep. 7 (2017) 10117, https://doi.org/10.1038/s41598-017-10371-5.
[26] G. Wang, L. He, C. Yuan, Y. Huang, Z. Liu, C. Liang, Pretreatment MR imaging radiomics signatures for response prediction to induction chemotherapy in patients with nasopharyngeal carcinoma, Eur. J. Radiol. 98 (2018) 100–106, https://doi.org/10.1016/j.ejrad.2017.11.007.
[27] B. Zhang, X. He, F. Ouyang, D. Gu, Y. Dong, L. Zhang, et al., Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma, Cancer Lett. 403 (2017) 21–27, https://doi.org/10.1016/j.canlet.2017.06.004.
[28] B. Zhang, F. Ouyang, D. Gu, Y. Dong, L. Zhang, X. Mo, et al., Advanced nasopharyngeal carcinoma: pre-treatment prediction of progression based on multi-parametric MRI radiomics, Oncotarget 8 (2017) 72457–72465. https://doi.org/10.18632/oncotarget.19799.
[29] S. Li, K. Wang, Z. Hou, J. Yang, W. Ren, S. Gao, 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. 8 (2018) 648, https://doi.org/10.3389/fonc.2018.00648.
[30] W. Lv, Q. Yuan, Q. Wang, J. Ma, J. Jiang, W. Yang, et al., Robustness versus disease differentiation when varying parameter settings in radiomics features: application to nasopharyngeal PET/CT, Eur. Radiol. 28 (2018) 3245–3254, https://doi.org/10.1007/s00330-018-5343-0.
[31] M. D. Anderson Cancer Center Head and Neck Quantitative Imaging Working Group, Elhalawani H, Kanwar A, Mohamed ASR, White A, Zafereo J, et al. Investigation of radiomic signatures for local recurrence using primary tumor texture analysis in oropharyngeal head and neck cancer patients. Sci Rep 2018;8: 1524. https://doi.org/10.1038/s41598-017-14687-0.
[32] J. Ren, J. Tian, Y. Yuan, D. Dong, X. Li, Y. Shi, et al., Magnetic resonance imaging based radiomics signature for the preoperative discrimination of stage I-II and III-IV head and neck squamous cell carcinoma, Eur. J. Radiol. 106 (2018) 1–6, https://doi.org/10.1016/j.ejrad.2018.07.002.
[33] L. Chen, Z. Zhou, D. Sher, Q. Zhang, J. Shah, N.-L. Pham, et al., Combining many-objective radiomics and 3D convolutional neural network through evidential reasoning to predict lymph node metastasis in head and neck cancer, Phys. Med. Biol. 64 (2019) 075011, https://doi.org/10.1088/1361-6560/ab083a.
[34] D. Du, H. Feng, W. Lv, S. Ashrafinia, Q. Yuan, Q. Wang, et al., Machine Learning Methods for Optimal Radiomics-Based Differentiation Between Recurrence and Inflammation: Application to Nasopharyngeal Carcinoma Post-therapy PET/CT Images, Mol. Imag. Biol. 22 (2020) 730–738, https://doi.org/10.1007/s11307-019-01411-9.
[35] R. Forghani, A. Chatterjee, C. Reinhold, A. Pérez-Lara, G. Romero-Sanchez, Y. Ueno, et al., Head and neck squamous cell carcinoma: prediction of cervical lymph node metastasis by dual-energy CT texture analysis with machine learning, Eur. Radiol. 29 (2019) 6172–6181, https://doi.org/10.1007/s00330-019-06159-y.
[36] Y. Liu, H. Shi, S. Huang, X. Chen, H. Zhou, H. Chang, et al., Early prediction of acute xerostomia during radiation therapy for nasopharyngeal cancer based on delta radiomics from CT images, Quant. Imaging Med. Surg. 9 (2019) 1288–1302. https://doi.org/10.21037/qims.2019.07.08.
[37] T. Liu, S. Zhou, J. Yu, Y. Guo, Y. Wang, J. Zhou, et al., Prediction of Lymph Node Metastasis in Patients With Papillary Thyroid Carcinoma: A Radiomics Method Based on Preoperative Ultrasound Images, Technol. Cancer Res. Treat. 18 (2019) 153303381983171, https://doi.org/10.1177/1533033819831713.
[38] X. Ming, R.W. Oei, R. Zhai, F. Kong, C. Du, C. Hu, et al., MRI-based radiomics signature is a quantitative prognostic biomarker for nasopharyngeal carcinoma, Sci. Rep. 9 (2019) 10412, https://doi.org/10.1038/s41598-019-46985-0.
[39] W. Wu, J. Ye, Q. Wang, J. Luo, S. Xu, CT-Based Radiomics Signature for the Preoperative Discrimination Between Head and Neck Squamous Cell Carcinoma Grades, Front. Oncol. 9 (2019) 821, https://doi.org/10.3389/fonc.2019.00821.
[40] T. Yu, S. Lam, L. To, K. Tse, N. Cheng, Y. Fan, et al., Pretreatment Prediction of Adaptive Radiation Therapy Eligibility Using MRI-Based Radiomics for Advanced Nasopharyngeal Carcinoma Patients, Front. Oncol. 9 (2019) 1050, https://doi.org/10.3389/fonc.2019.01050.
[41] L. Zhang, D. Dong, H. Li, J. Tian, F. Ouyang, X. Mo, et al., Development and validation of a magnetic resonance imaging-based model for the prediction of distant metastasis before initial treatment of nasopharyngeal carcinoma: A

retrospective cohort study, EBioMedicine 40 (2019) 327–335, https://doi.org/10.1016/j.ebiom.2019.01.013.
[42] E.-H. Zhuo, W.-J. Zhang, H.-J. Li, G.-Y. Zhang, B.-Z. Jing, J. Zhou, et al., Radiomics on multi-modalities MR sequences can subtype patients with non-metastatic nasopharyngeal carcinoma (NPC) into distinct survival subgroups, Eur. Radiol. 29 (2019) 5590–5599, https://doi.org/10.1007/s00330-019-06075-1.
[43] F. Akram, P.E. Koh, F. Wang, S. Zhou, S.H. Tan, M. Paknezhad, et al., Exploring MRI based radiomics analysis of intratumoral spatial heterogeneity in locally advanced nasopharyngeal carcinoma treated with intensity modulated radiotherapy, PLoS One 15 (2020) e0240043.
[44] Q. Feng, J. Liang, L. Wang, J. Niu, X. Ge, P. Pang, et al., Radiomics Analysis and Correlation With Metabolic Parameters in Nasopharyngeal Carcinoma Based on PET/MR Imaging, Front. Oncol. 10 (2020) 1619, https://doi.org/10.3389/fonc.2020.01619.
[45] R. Guo, J. Guo, L. Zhang, X. Qu, S. Dai, R. Peng, et al., CT-based radiomics features in the prediction of thyroid cartilage invasion from laryngeal and hypopharyngeal squamous cell carcinoma, Cancer Imaging 20 (2020) 81, https://doi.org/10.1186/s40644-020-00359-2.
[46] Haider SP, Zeevi T, Baumeister P, Reichel C, Sharaf K, Forghani R, et al. Potential Added Value of PET/CT Radiomics for Survival Prognostication beyond AJCC 8th Edition Staging in Oropharyngeal Squamous Cell Carcinoma. Cancers 2020;12: 1778. https://doi.org/10.3390/cancers12071778.
[47] S.P. Haider, A. Mahajan, T. Zeevi, P. Baumeister, C. Reichel, K. Sharaf, et al., PET/CT radiomics signature of human papilloma virus association in oropharyngeal squamous cell carcinoma, Eur. J. Nucl. Med. Mol. Imaging 47 (2020) 2978–2991, https://doi.org/10.1007/s00259-020-04839-2.
[48] T.-Y. Ho, C.-H. Chao, S.-C. Chin, S.-H. Ng, C.-J. Kang, N.-M. Tsang, Classifying Neck Lymph Nodes of Head and Neck Squamous Cell Carcinoma in MRI Images with Radiomic Features, J. Digit. Imaging 33 (2020) 613–618, https://doi.org/10.1007/s10278-019-00309-w.
[49] F. Li, D. Pan, Y. He, Y. Wu, J. Peng, J. Li, et al., Using ultrasound features and radiomics analysis to predict lymph node metastasis in patients with thyroid cancer, BMC Surg. 20 (2020) 315, https://doi.org/10.1186/s12893-020-00974-7.
[50] Mukherjee P, Cintra M, Huang C, Zhou M, Zhu S, Colevas AD, et al. CT-based Radiomic Signatures for Predicting Histopathologic Features in Head and Neck Squamous Cell Carcinoma. Radiology: Imaging Cancer 2020;2:e190039. https://doi.org/10.1148/rycan.2020190039.
[51] V. Romeo, R. Cuocolo, C. Ricciardi, L. Ugga, S. Cocozza, F. Verde, et al., Prediction of Tumor Grade and Nodal Status in Oropharyngeal and Oral Cavity Squamous-cell Carcinoma Using a Radiomic Approach, Anticancer Res 40 (2020) 271–280. https://doi.org/10.21873/anticanres.13949.
[52] W.T. Tran, H. Suraweera, K. Quiaoit, D. DiCenzo, K. Fatima, D. Jang, et al., Quantitative ultrasound delta-radiomics during radiotherapy for monitoring treatment responses in head and neck malignancies, Future Sci. OA (2020;6: FSO624.), https://doi.org/10.2144/fsoa-2020-0073.
[53] B. Zhang, Z. Lian, L. Zhong, X. Zhang, Y. Dong, Q. Chen, et al., Machine-learning based MRI radiomics models for early detection of radiation-induced brain injury in nasopharyngeal carcinoma, BMC Cancer 20 (2020) 502, https://doi.org/10.1186/s12885-020-06957-4.
[54] K. Fatima, A. Dasgupta, D. DiCenzo, C. Kolios, K. Quiaoit, M. Saifuddin, et al., Ultrasound delta-radiomics during radiotherapy to predict recurrence in patients with head and neck squamous cell carcinoma, Clinical and Translational Radiation Oncology 28 (2021) 62–70, https://doi.org/10.1016/j.ctro.2021.03.002.
[55] S.P. Haider, K. Sharaf, T. Zeevi, P. Baumeister, C. Reichel, R. Forghani, et al., Prediction of post-radiotherapy locoregional progression in HPV-associated oropharyngeal squamous cell carcinoma using machine-learning analysis of baseline PET/CT radiomics, Transl. Oncol. 14 (2021) 100906, https://doi.org/10.1016/j.tranon.2020.100906.
[56] M.-J. Kim, Y. Choi, Y.E. Sung, Y.S. Lee, Y.-S. Kim, K.-J. Ahn, et al., Early risk-assessment of patients with nasopharyngeal carcinoma: the added prognostic value of MR-based radiomics, Transl. Oncol. 14 (2021) 101180, https://doi.org/10.1016/j.tranon.2021.101180.
[57] L. Peng, X. Hong, Q. Yuan, L. Lu, Q. Wang, W. Chen, Prediction of local recurrence and distant metastasis using radiomics analysis of pretreatment nasopharyngeal [18F]FDG PET/CT images, Ann. Nucl. Med. 35 (2021) 458–468, https://doi.org/10.1007/s12149-021-01585-9.
[58] F. Wang, R. Tan, K. Feng, J. Hu, Z. Zhuang, C. Wang, et al., Magnetic Resonance Imaging-Based Radiomics Features Associated with Depth of Invasion Predicted Lymph Node Metastasis and Prognosis in Tongue Cancer, Magn. Reson. Imaging 56 (2022) 196–209, https://doi.org/10.1002/jmri.28019.
[59] S. Woolen, A. Virkud, L. Hadjiiski, K. Cha, H.-P. Chan, P. Swiecicki, et al., Prediction of Disease Free Survival in Laryngeal and Hypopharyngeal Cancers Using CT Perfusion and Radiomic Features: A Pilot Study, Tomography 7 (2021) 10–19, https://doi.org/10.3390/tomography7010002.
[60] Y. Yang, M. Wang, K. Qiu, Y. Wang, X. Ma, Computed tomography-based deep-learning prediction of induction chemotherapy treatment response in locally advanced nasopharyngeal carcinoma, Strahlenther. Onkol. 198 (2022) 183–193, https://doi.org/10.1007/s00066-021-01874-2.
[61] M.H. Zhang, D. Cao, D.T. Ginat, Radiomic Model Predicts Lymph Node Response to Induction Chemotherapy in Locally Advanced Head and Neck Cancer, Diagnostics 11 (2021) 588, https://doi.org/10.3390/diagnostics11040588.
[62] J. Zhong, R. Frood, P. Brown, H. Nelstrop, R. Prestwich, G. McDermott, et al., Machine learning-based FDG PET-CT radiomics for outcome prediction in larynx and hypopharynx squamous cell carcinoma, Clin. Radiol. 76 (2021) 78.e9–78.e17, https://doi.org/10.1016/j.crad.2020.08.030.
[63] J. Kaźmierska, M.R. Kaźmierski, T. Bajon, T. Winiecki, A. Bandurska-Luque, A. Ryczkowski, et al., Prediction of Incomplete Response of Primary Tumour Based on Clinical and Radiomics Features in Inoperable Head and Neck Cancers after Definitive Treatment, JPM 12 (2022) 1092, https://doi.org/10.3390/jpm12071092.
[64] M. Kim, J.H. Lee, L. Joo, B. Jeong, S. Kim, S. Ham, et al., Development and Validation of a Model Using Radiomics Features from an Apparent Diffusion Coefficient Map to Diagnose Local Tumor Recurrence in Patients Treated for Head and Neck Squamous Cell Carcinoma, Korean J. Radiol. 23 (2022) 1078, https://doi.org/10.3348/kjr.2022.0299.
[65] M. Nakajo, K. Kawaji, H. Nagano, M. Jinguji, A. Mukai, H. Kawabata, et al., The Usefulness of Machine Learning-Based Evaluation of Clinical and Pretreatment [18F]-FDG-PET/CT Radiomic Features for Predicting Prognosis in Hypopharyngeal Cancer, Mol. Imag. Biol. 25 (2023) 303–313, https://doi.org/10.1007/s11307-022-01757-7.
[66] Y. Xi, X. Ge, H. Ji, L. Wang, S. Duan, H. Chen, et al., Prediction of Response to Induction Chemotherapy Plus Concurrent Chemoradiotherapy for Nasopharyngeal Carcinoma Based on MRI Radiomics and Delta Radiomics: A Two-Center Retrospective Study, Front. Oncol. 12 (2022) 824509, https://doi.org/10.3389/fonc.2022.824509.
[67] M. Lin, X. Tang, L. Cao, Y. Liao, Y. Zhang, J. Zhou, Using ultrasound radiomics analysis to diagnose cervical lymph node metastasis in patients with nasopharyngeal carcinoma, Eur. Radiol. 33 (2022) 774–783, https://doi.org/10.1007/s00330-022-09122-6.
[68] Y.-P. Zhang, X.-Y. Zhang, Y.-T. Cheng, B. Li, X.-Z. Teng, J. Zhang, et al., Artificial intelligence-driven radiomics study in cancer: the role of feature engineering and modeling, Military Med Res 10 (2023) 22, https://doi.org/10.1186/s40779-023-00458-8.
[69] F. Wang, B. Zhang, X. Wu, L. Liu, J. Fang, Q. Chen, et al., Radiomic Nomogram Improves Preoperative T Category Accuracy in Locally Advanced Laryngeal Carcinoma, Front. Oncol. 9 (2019) 1064, https://doi.org/10.3389/fonc.2019.01064.
[70] M. Bologna, V. Corino, G. Calareso, C. Tenconi, S. Alfieri, N.A. Iacovelli, et al., Baseline MRI-Radiomics Can Predict Overall Survival in Non-Endemic EBV-Related Nasopharyngeal Carcinoma Patients, Cancers 12 (2020) 2958, https://doi.org/10.3390/cancers12102958.
[71] K. Liu, Q. Qiu, Y. Qin, T. Chen, D. Zhang, L. Huang, et al., Radiomics Nomogram Based on Multiple-Sequence Magnetic Resonance Imaging Predicts Long-Term Survival in Patients Diagnosed With Nasopharyngeal Carcinoma, Front. Oncol. 12 (2022) 852348, https://doi.org/10.3389/fonc.2022.852348.
[72] Z. Zhou, L. Chen, D. Sher, Q. Zhang, J. Shah, N.-L. Pham, et al., in: Predicting Lymph Node Metastasis in Head and Neck Cancer by Combining Many-Objective Radiomics and 3-Dimensioal Convolutional Neural Network through Evidential Reasoning, IEEE, Honolulu, HI, USA, 2018, pp. 1–4, https://doi.org/10.1109/EMBC.2018.8513070.
[73] X. Zhong, L. Li, H. Jiang, J. Yin, B. Lu, W. Han, et al., Cervical spine osteoradionecrosis or bone metastasis after radiotherapy for nasopharyngeal carcinoma? The MRI-based radiomics for characterization, BMC Med. Imaging 20 (2020) 104, https://doi.org/10.1186/s12880-020-00502-2.
[74] A.D. King, J.F. Griffith, J.M. Abrigo, S. Leung, F. Yau, G.M.K. Tse, et al., Osteoradionecrosis of the upper cervical spine: MR imaging following radiotherapy for nasopharyngeal carcinoma, Eur. J. Radiol. 73 (2010) 629–635, https://doi.org/10.1016/j.ejrad.2008.12.016.
[75] R. Forghani, P. Savadjiev, A. Chatterjee, N. Muthukrishnan, C. Reinhold, B. Forghani, Radiomics and Artificial Intelligence for Biomarker and Prediction Model Development in Oncology, Comput. Struct. Biotechnol. J. 17 (2019) 995–1008, https://doi.org/10.1016/j.csbj.2019.07.001.
[76] R.T.H.M. Larue, G. Defraene, D. De Ruysscher, P. Lambin, W. Van Elmpt, Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures, BJR 90 (2017) 20160665, https://doi.org/10.1259/bjr.20160665.
[77] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: Images Are More than Pictures, They Are Data. Radiology 278 (2016) 563–577, https://doi.org/10.1148/radiol.2015151169.
[78] M. Avanzo, L. Wei, J. Stancanello, M. Vallières, A. Rao, O. Morin, et al., Machine and deep learning methods for radiomics, Med. Phys. 47 (2020), https://doi.org/10.1002/mp.13678.
[79] R.O. Alabi, A. Sjöblom, T. Carpén, M. Elmusrati, I. Leivo, A. Almangush, et al., Application of artificial intelligence for overall survival risk stratification in oropharyngeal carcinoma: A validation of ProgTOOL, Int. J. Med. Inf. 175 (2023) 105064, https://doi.org/10.1016/j.ijmedinf.2023.105064.
[80] R.O. Alabi, O. Youssef, M. Pirinen, M. Elmusrati, A.A. Mäkitie, I. Leivo, et al., Machine learning in oral squamous cell carcinoma: Current status, clinical concerns and prospects for future—A systematic review, Artif. Intell. Med. 115 (2021) 102060, https://doi.org/10.1016/j.artmed.2021.102060.
[81] S.-K. Lam Y. Zhang B. Li J.-C. Sun C.-Y.-T. Liu et al. Multi-Organ Omics-Based Prediction for Adaptive Radiation Therapy Eligibility in Nasopharyngeal Carcinoma Patients Undergoing Concurrent Chemoradiotherapy. Front Oncol 2022;11:792024. 10.3389/fonc.2021.792024.
[82] Y. Zhang, K. Xia, Y. Jiang, P. Qian, W. Cai, C. Qiu, et al., Multi-modality Fusion & Inductive Knowledge Transfer Underlying Non-Sparse Multi-Kernel Learning and Distribution Adaption, IEEE/ACM Trans Comput Biol and Bioinf (2022) 1, https://doi.org/10.1109/TCBB.2022.3142748.
[83] W. Li, H. Shen, L. Han, J. Liu, B. Xiao, X. Li, et al., A Multiparametric Fusion Radiomics Signature Based on Contrast-Enhanced MRI for Predicting Early

Recurrence of Hepatocellular Carcinoma, J. Oncol. 2022 (2022) 1–12, https://doi.org/10.1155/2022/3704987.

[84] S.A. Keek, F.W.R. Wesseling, H.C. Woodruff, J.E. Van Timmeren, I.H. Nauta, T. K. Hoffmann, et al., A Prospectively Validated Prognostic Model for Patients with Locally Advanced Squamous Cell Carcinoma of the Head and Neck Based on Radiomics of Computed Tomography Images, Cancers 13 (2021) 3271, https://doi.org/10.3390/cancers13133271.

[85] K. Sheikh, S.H. Lee, Z. Cheng, P. Lakshminarayanan, L. Peng, P. Han, et al., Predicting acute radiation induced xerostomia in head and neck Cancer using MR and CT Radiomics of parotid and submandibular glands, Radiat. Oncol. 14 (2019) 131, https://doi.org/10.1186/s13014-019-1339-4.

[86] C. Pan, P. Zhou, J. Tan, B. Sun, R. Guan, Z. Wang, et al., in: Liver Tumor Detection via A Multi-Scale Intermediate Multi-Modal Fusion Network on MRI Images, IEEE, Anchorage, AK, USA, 2021, pp. 299–303, https://doi.org/10.1109/ICIP42928.2021.9506237.

[87] M. Bogowicz, A. Jochems, T.M. Deist, S. Tanadini-Lang, S.H. Huang, B. Chan, et al., Privacy-preserving distributed learning of radiomics to predict overall survival and HPV status in head and neck cancer, Sci. Rep. 10 (2020) 4542, https://doi.org/10.1038/s41598-020-61297-4.

[88] R.O. Alabi, M. Elmusrati, I. Leivo, A. Almangush, A.A. Mäkitie, Machine learning explainability in nasopharyngeal cancer survival using LIME and SHAP, Sci. Rep. 13 (2023) 8984, https://doi.org/10.1038/s41598-023-35795-0.