The prognostic value of CT radiomic features from primary tumours and pathological lymph nodes in head and neck cancer patients

Zhai, Tiantian

DOI:
10.33612/diss.111448998

CHAPTER 6

Discussion and future persepectives

6.1 Summary and general discussion

Large quantities of medical images are generated every day. At present, in oncology, these are mainly used as a qualitative tool for diagnostic and staging purposes. Tumour phenotypic differences in patients can be visualized by imaging. However, images may contain more than meets the eye [1,2]. For example, they can also contain valuable unexplored information that is not visible to the naked eye. Novel computational technologies radiomics, which was first proposed in 2012, can help to convert qualitative image information into high-dimensional quantitative and mineable data [3,4]. These so called radiomic features or image biomarkers represent the intensity, shape and textural heterogeneity of a region of interest (ROI), which have a great potential for improving treatment outcome predictions [3,5]. The overall purpose of this thesis is to provide better understanding of radiomic features in predicting treatment failures of head and neck cancer (HNC) patients after definitive non-surgical treatment and to investigate their additional prognostic power compared to classical prognostic factors.

6.1.1 Association of radiomic data with treatment failures

Before 2017, the potential prognostic value of radiomic features, significantly associated with survival in HNC, has been demonstrated. Aerts et al. identified the best performing radiomic features from four radiomic categories (intensity, geometry/shape, texture and wavelet decomposition) to create a signature: statistics energy, shape compactness, gray level non-uniformity, and wavelet (HLH) gray level non-uniformity with a concordance index (c-index) of 0.69 in predicting the overall survival (OS) of HNC [3]. This radiomic signature was externally validated by Leijenaar et al. in a subsequent study in a large North American cohort of oropharyngeal squamous cell carcinoma with c-indexes of 0.63-0.65 [6]. In 2017, we confirmed the prognostic value of radiomics for predicting OS using 289 nasopharyngeal cancer patients from Shantou University Medical College (SUMC), China and 298 HNC patients from the University Medical Centre of Groningen (UMCG), the Netherlands, which are presented in **Chapter 2.** In contrast to Aerts' model, we included the most frequently selected radiomic features from 1000 bootstrap samples instead of combining the four-best performing radiomic features from each radiomic category. In this way, we could reduce the probability of over-fitting [7,8]. Our study identified a geometric feature (volume density) and a texture feature (run length

non-uniformity) of the primary tumour as independent prognostic factors. Moreover, we were the first group that included radiomic features of the pathologic lymph nodes (pLN) and reported major-axis-length of pLN as a significant prognostic factor (Table 1). The multivariable model for OS based on the three radiomic features resulted in a

Table 1 Prognostic CT-based radiomic features (image-biomarkers) for overall survival (OS), local control (LC), regional control (RC), distant metastasis-free survival (DMFS), disease-free survival (DFS), and nodal control (NC) rate in head and neck cancer (HNC) patients.

Chapter No.	Endpoint	No. of Observations (training/validation)	Anatomy site (training/validation)	Prognostic radiomic features	ROI	Radiomic Category
2	OS	289/298	NPC/HNC	Run Length Non-uniformity	Primary tumour	Texture: GLRLM
				Volume-density	Primary tumour	Geometry
				Major-axis-length	Lymph node	Geometry
3	LC	240/204	HNC/HNC	Correlation	Primary tumour	Texture: GLCM
	RC	240/204	HNC/HNC	Bounding-box-volume	Primary tumour	Geometry
				Major-axis-length	Lymph node	Geometry
	DMFS	240/204	HNC/HNC	Bounding-box-volume	Primary tumour	Geometry
				Major-axis-length	Lymph node	Geometry
	DFS	240/204	HNC/HNC	Bounding-box-volume	Primary tumour	Geometry
				Correlation	Lymph node	Texture: GLCM
4	NC	558/467	HNC	Correlation	Lymph node	Texture: GLCM
				Least-axis-length	Lymph node	Geometry
	NC	268/274	OPC	Short run high grey level emphasis	Lymph node	Texture: GLRLM
				Least-axis-length	Lymph node	Geometry

Abbreviations: NPC = nasopharyngeal caner; OPC = oropharyngeal cancer; ROI = region of interest; GLRLM = gray level run-length matrix; GLCM = gray level co-occurrence matrix.

c-index of 0.72 in SUMC training cohort and 0.67 in the UMCG validation cohort. We also showed that the radiomic features improved the prognostic performance of the models containing clinical factors significantly, which was not investigated in the study of Aerts et al. The additional prognostic value of radiomic data was discussed extensively in 6.1.3. Subsequently, we performed a detailed analysis on the different patterns of failure by evaluating local control (LC), regional control (RC), distant metastasis-free survival (DMFS) and disease-free survival (DFS) in a large cohort of 444 HNC patients from UMCG described in **Chapter 3**. The patient cohort was split into a model development and validation cohort of 240 and 204 patients, respectively. The following radiomic features were most frequently selected and significantly associated with the endpoints in the multivariable analyses for LC: correlation of gray level co-occurrence matrix (GLCM); for RC and DMFS: bounding-box-volume and major-axis-length of pLN, and for DFS: bounding-box-volume and correlation of GLCM. More detailed information is given in Table 1. With only one or two radiomic features, the LC, RC, DMFS, and DFS models' performances in the validation cohort were 0.62, 0.80. 0.68, and 0.65 respectively. The endpoints reflecting the specific pattern of relapse may provide more valuable information than survival prediction in guiding future treatment intensification targeted on patients with specific high risk of loco-regional failure or distant metastasis.

Approximately 65% of the HNC patients present with pathological lymph nodes, and more than 80% of them have more than one pathological lymph node [9]. From clinical practice we know that different lymph nodes within one HNC patients may show heterogeneity with regard to response to non-surgical treatment. This phenomenon might open new avenues to use prediction models to support clinical decision making, not only for individual patients, but also for specific lymph nodes within one patient. Therefore, we explored the association between radiomic features from individual pathological lymph nodes and nodal failures in 277 node-positive patients with 1,025 pathological lymph nodes (pLN) as described in **Chapter 4.** We found that lymph nodes with large values of least-axis-length and correlation of GLCM showed worse nodal control. When tested in the validation cohort as well as subgroup cohort of oropharyngeal cancer, the model with two radiomic features preserved good discrimination with c-indexes of 0.79 and 0.86, respectively. The model was further externally validated in a TRIPOD type 4 modelling

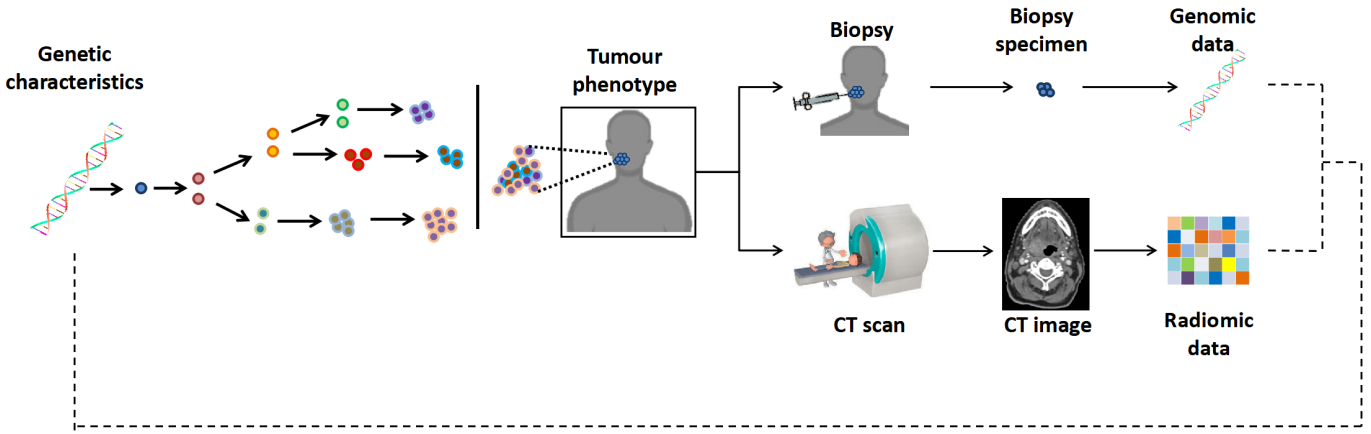

Figure 1. The relationships between genomic data and radiomic data. This figure shows the whole process of tumour growth and the analysis process. Genetic mutations cause the multiple subclonal populations coexisting in the tumour. Multiple subclonal populations grow into macroscopic tumour over time. Genomic data is assessed using the biopsy specimens. Radiomic data is extracted from the three-dimensional contoured tumour based on CT image, which can provide the complementary information to genomic data based on biopsy specimens.

study. This study consisted of 113 patients with 374 pathological lymph nodes treated at Maastro, Maastricht, the Netherlands. The model showed good discrimination with a c-index of 0.71 and good agreement between predicted and observed nodal control probabilities (**Chapter 5**). External validation in an independent cohort is an essential step before the model can be used in the clinic to support treatment decisions for head and neck squamous cell carcinoma (HNSCC) patients. A discussion on the next steps into clinical applications will be given in the "future perspective" of this chapter.

6.1.2 Association of radiomic data with tumour biology

The role of radiomic features in predicting treatment outcomes in patients with HNSCC is promising. Radiomic features provide quantitative information on tumour phenotypes at a macroscopic level, which is a result of genetic and microenvironmental characteristics [10,11]. In other words, it can be expected that correlations exist between radiomic features and genomic data (Figure 1).

Aerts et al. compared the radiomic features with gene-expression profiles and showed that radiomic features from lung cancer significantly correlated with different biologic gene sets. It is worth mentioning that the textural radiomic features strongly correlated with cell cycling pathways which regulate cell proliferation and evolution [3]. Another study, based on a small cohort of 27 oral cavity squamous carcinoma patients, suggested that the expression of vascular endothelial growth factor (VEGF) receptors 1 and 2 significantly correlated with the enhancement (textural features) of the tumour, while cyclin D1 expression and the regulator of apoptosis highly correlated with the tumour size (geometric features) [12,13]. Recently, Zhu et al. performed a comprehensive study to explore the relationship between the tumour genomic system and the multiple aspects of tumour image features for HNSCC. They identified wide-spread and statistically significant associations between various genomic features and radiomic features characterizing the size, shape, and texture of the tumour. In this study, more transcriptional activities were found in patients with radiomic features representing larger, more irregular and heterogeneous tumours. DNA methylation changes were negatively associated with tumour size features, while miRNA over-expression was positively associated with tumour texture homogeneity [14].

There are two main types of radiomic features reported in these radiomics-genomics

association studies, namely geometric and textural features. This is in line with the findings as presented in this thesis. Four geometric features (volume-density, major-axis-length, bounding-box-volume, and least-axis-length) showed significant associations with all endpoints, and three textural features (run length non-uniformity, correlation of GLCM, and short run high gray level emphasis of gray level run-length matrix (GLRLM)) showed significant association with OS, LC, DFS, and NC.

Geometric features may contain similar information as classical clinical features, such as tumour volume, tumour diameters. Moreover, in this thesis we showed that geometric features can represent more complex geometric phenotypes, such as bounding-box-volume which not only gives information on the tumour volume, but also in the irregularity of the tumour shape. Bounding-box-volume, as reported in **Chapter 3**, which refers to the volume of the smallest cube that encloses all pixels of the contoured tumour, was more predictive for RC, DMFS and DFS than the classical clinical feature tumour volume. Some examples are shown in Fig. 4a and 4b in Chapter 3. The irregular tumour in Fig. 4b had a similar tumour volume as the tumour in Fig. 4a, but the bounding-box-volume of Fig. 4b was twice as large as that of Fig. 4a. In our opinion, tumour volume depends on the doubling time of tumour cells, while tumour shape is more determined by invasive growth patterns. Therefore geometric radiomic features that represent the tumour shape could be related to the growth patterns and aggressive behaviours which are actually driven by gene mutations such as inactivation of tumour suppressor genes, dysfunction of regulators of apoptosis, over expression of VEGF, etc. [15] We expect that specific geometric features may be useful substitutes for classical clinical features to improve prediction of treatment failure in the future.

Another important type of radiomic features is the textural features. They decode the tumour radiological heterogeneity by quantifying the presence of different patterns in voxel intensities in the three-dimensional contoured tumour. The textural feature Correlation of GLCM (Corre-GLCM) describes the correlation of a reference voxel to its neighbours. Tumours with lower values of Corre-GLCM have large areas of similar intensities, i.e. lower radiological heterogeneity, which was associated with lower local failure and nodal failure rates in **Chapter 3, 4 and 5**. Visual inspection of the CT images with higher values of Corre-GLCM suggested that the presence of necrosis, which can be

recognized as an area with lower CT-intensities surrounded by an irregular rim of higher CT-intensities in contrast-enhanced CT images, increased the value of Corre-CLCM [16]. Therefore, in our opinion, Corre-GLCM may indicate the necrosis status of the tumour. The textural feature short run high gray emphasis (SRHGE) emphasises small areas with high CT-intensities. In this thesis, a lower SRHGE was associated with a higher nodal failure risk. Lower SRHGE values are expected in volumes with lower contrast enhancement (all included CT-scans were contrast enhanced). Since contrast enhancement is related to local circulation, lower SRHGE values can be expected in hypo-vascular volumes with a higher risk of hypoxia [17,18]. Accordingly, higher failure risk in lymph nodes with a lower SRHGE could be associated with the oxygenation status of the lymph node.

Although we did not evaluate direct evidence on the relation between gene expressions, microenvironment and radiomic features, the radiomic features identified in this thesis are related to different tumour phenotypes such as tumour shape and necrosis (Figure 1). From a biological point of view, the different tumour phenotypes are caused by a series of genetic mutations and heterogeneity of the microenvironment. For example, tumour necrosis is associated with multiple gene sets of IL6, CXCL8, SERPINE1 related to hypoxia, angiogenesis and inflammation [19]; irregular shape can be associated with DNA methylation changes, dysfunction of regulators of apoptosis, and more transcriptional activities [14]. A series of genetic mutations causes multiple subclonal populations coexisting in the tumour and regulates the proliferation and evolution of the tumour cells. Subsequently, differences in growth speed and pattern between the subclonal populations in combination with microenvironment heterogeneity result in tumour phenotypic heterogeneity [20–23] (Figure 1). The evidence on the associations between radiomic features and tumour heterogeneity related gene sets is not yet clear. Further research on large dataset is necessary to confirm the associations between radiomics and genomics, and develop a theoretical basis of biology for these associations [20,24]. We expect that radiomics could become a non-invasive tool that can provide data on tumour phenotypes which is complementary to data from limited biopsy samples.

6.1.3 Additional prognostic value of radiomic data

Several studies have shown the diagnostic, prognostic and predictive value of radiomic features. Furthermore, association analyses have revealed a possible relation between

radiomic features and genetic phenotypes. However, the extent to which radiomic features can provide additional prognostic information to current clinical features in predicting HNC patient treatment failures, is less clear. Therefore, the first step in this thesis was to build optimal prediction models consisting of classical prognostic factors only as reference models (clinical models). Additionally, we developed models containing quantitative radiomic features only (radiomic models). In **Chapter 2 and 3**, we were the first to prove that radiomic models containing quantitative radiomic features describing the volume, irregular shape and radiological heterogeneity of the tumour and the distance between lymph nodes can perform as well as clinical variables in predicting LC, RC, DMFS, DFS and OS for HNC patients. However, when clinical and radiomics models were combined, model performance remained similar or became slightly better than clinical models in predicting LC, RC and DMFS. For the prediction of DFS and OS, the combined models performed significantly better than the clinical model. Patients stratified with the combined models showed greater differences between the low and high-risk groups in the validation cohort than with clinical models for LC, RC, DFS and OS (Figure 3 in Chapter 3). These combined models may be used to identify patients with high risks of recurrence and survival prior to treatment. After we published our results, Vallieres et al. confirmed our findings and demonstrated that the prediction models combining radiomic and clinical variables yielded a better performance than models with clinical variables or radiomic features only [25].

In **Chapter 4 and 5**, clinical, radiomic and combined models for individual lymph node failure prediction were built and compared. The c-index of the radiomic model was 0.84, slightly but not significantly ($p = 0.093$) better than that of the clinical model 0.78 in the UMCG training cohort. When tested in the UMCG validation cohort, the c-index of the radiomic model was 0.79 and was higher than that (0.69) of the clinical model. Further validation in the external Maastro cohort showed that the c-index of the clinical model was significantly worse (0.57) than that of the radiomics model (0.71) in this cohort. This illustrates that the clinical model based on patient-specific features may be less robust than the radiomic model based on individual lymph nodes for nodal failure prediction. Moreover, the radiomic features extracted from each lymph node are expected to be more informative than clinical features as they can be measured objectively and better

quantified than the categorical clinical features such as N-stage and presence or absence of central necrosis. This could be the reason why radiomic features outperform clinical variables in predicting individual lymph failures.

When radiomic features were combined with clinical variables, the c-index significantly improved from 0.78 to 0.90 ($p < 0.001$). With the combined model, better discrimination and calibration was obtained in the UMCG as well as in the Maastro validation cohorts. The combined model enables a more accurate risk stratification of pathological lymph nodes and provide opportunities for more personalized decision making targeted on the individual tumour lesions [26–28].

To conclude, prediction models containing clinical factors and radiomic features perform as well or slightly better than models with clinical variables only for the prediction of LC, RC, DMFS, DFS and OS. Therefore, the additional prognostic values are limited for the treatment outcome prediction of individual patient. However, radiomic features of individual lymph nodes provide additional prognostic information regarding nodal failures.

6.2 Future perspectives

According to Robert J. Gillies, "Images are more than pictures, they are data" [2]. Progress in medical image analysis, especially the application of artificial intelligence, converts images into minable data. The efforts on the correlations between these data and the prognosis of cancer build the bridge between medical imaging and personalized medicine [26]. One could envision that, in the near future, radiomics, either combined with current classical prognostic features or not, may be gradually introduced into the clinic to guide a more personalized approach. We reviewed recent developments in radiomics including multiple imaging modalities, deep learning features, and the quality of data and models to propose schemes for future model improvement. Furthermore, we will discuss a set of clinical studies to validate the models described in this thesis. After a successful clinical validation, these models could be introduced into the clinic.

6.2.1 Future improvement

Multi-modality medical images

Radiomic features were extracted from computed tomography (CT) images in this

thesis. Data acquisition can be improved by including multi-modality medical images such as magnetic resonance imaging (MRI), positron emission tomography (PET), ultrasonography and in-treatment images such as weekly CT images.

CT is the most straightforward method to visualize the ROI variations in density and the easiest technique to compare across institutions due to the standard phantoms and clinical settings [29]. CT scans are always acquired in patients primarily treated with radiotherapy. Consequently, large datasets are available for analysis. Therefore, in most radiomic studies, CT images have been used [3,4,6].

As previously mentioned, genomic variations may affect cell proliferation, hypoxia or necrotic areas, which could translate into tumour metabolism heterogeneity. Therefore, functional PET may play an important role in radiomic analysis. Vallieres et al. found that in HNC patients, more radiomic features could be identified for LC and DMFS with 18F-FDG-PET than with CT images. However, models based on PET radiomic features failed to show better performance than CT radiomic features [25]. In addition, Bogowicz et al. compared PET and CT radiomic features for the prediction of LC, and confirmed that CT and PET radiomic features had equally good discriminative power[30].

CT radiomic features are calculated based on signal intensity (gray levels), generated from ROI density. In contrast to CT images, more sequences such as T1-weighted, T2-weighted, diffusion weighted imaging (DWI) provide better contrast resolution of the soft tissue which makes MR superior in identifying lesions. Therefore, radiomic analysis based on MR might be more accurate in identifying aggressive tumours than radiomics based on CT. For example, MRI is superior in detecting nasopharyngeal carcinoma than CT [31]. Zhang et al. identified radiomic features extracted from MRI as significant prognostic factors for advanced nasopharyngeal carcinoma with a better performance than what we found iin Chapter 2 [32,33]. Until now, studies based on MRI have been scarce, as large variations across institutions exist in acquisition settings and intensity standardisation, which have an impact on the robustness of radiomic features [28,34].

A few studies showed that ultrasound features may be used to distinguish malignant from benign tumours in the breast and thyroid [35,36]. So far, no studies in HNC have been published.

The reason for the limited use of MRI, PET and US is the lack of standard imaging protocols,

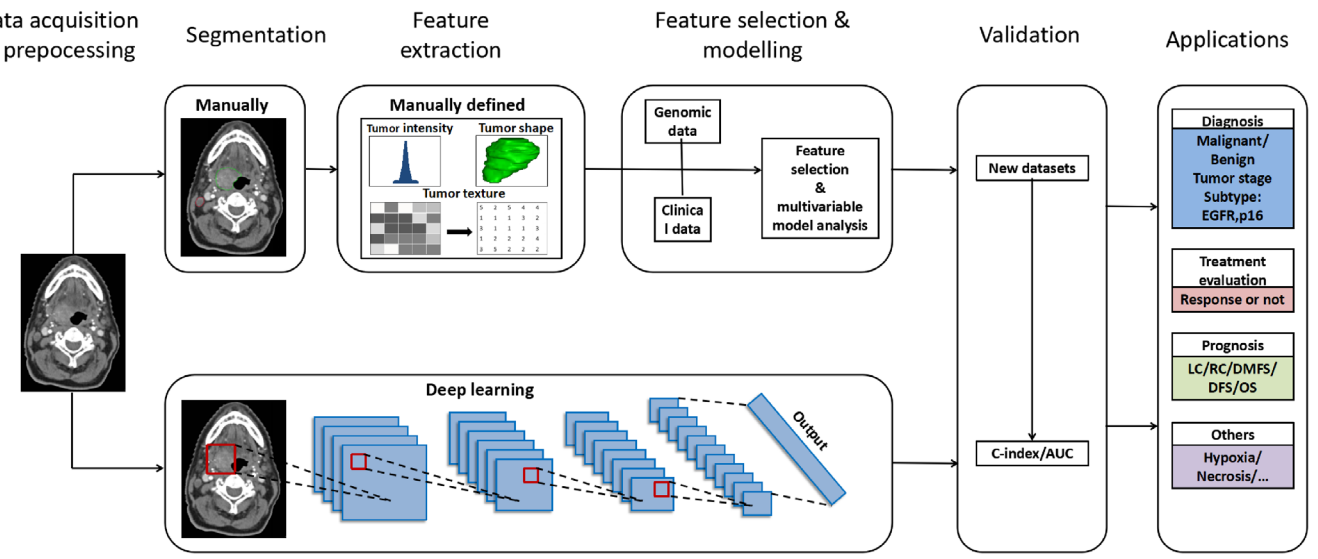

Figure 2. The radiomics pipeline of modelling and application with manually defined features and deep learning. It includes the main steps: data acquisition and pre-processing, tumour segmentation, feature extraction, feature selection and modelling, model validation and applications. Deep learning can be used for segmentation, feature extraction and selection, and modelling. External validation is the first step in making these models available for supporting future applications. Abbreviations: C-index = concordance index, AUC = area under the receiver operating characteristic curve, EGFR = epidermal growth factor receptor, LC = local control, RC = regional control, DMFS = distant metastasis free survival, DFS = disease free survival, OS = overall survival.

which has a large influence on the reproducibility and stability of quantitative features. In order to improve the applicability of different modalities for radiomic research, at least image acquisition and processing parameters should be reported in detail to allow the comparison between radiomic features from different institutions. Preferably, to enhance wide spread clinical application of prediction models including radiomic features, multicentre studies collecting large image datasets should use standardized acquisition and processing parameters and allow data sharing. [37,38].

Imaging during treatment may be another direction to improve current radiomic risk models. A recent study from Zwanenburg et al. showed that combining radiomic features from pre- and during-treatment CT images is a promising way to improve the loco-regional tumour control and OS prediction in HNC [39]. Wu et al. further confirmed the prognostic value of pre- and mid-treatment CT radiomics for predicting DMFS in oropharyngeal cancer patients [40]. Tumour changes, such as shrinkage and density change, reflect the response to the therapy, and therefore may provide additional prognostic value.

Deep learning in radiomics

An overview of the current radiomics workflow in this thesis is shown in Figure 2. We expanded this figure by adding deep learning features in head and neck cancer applications. Deep learning can be used for segmentation, feature extraction and selection, and classification [41–43].

That manual delineation is so time-consuming and labour-intensive currently limits the applicability of radiomic analysis in big data sets. Auto-segmentation using deep learning has the ability to reduce inter-observer variability and time with respect to manual delineation while considerable accuracy can be achieved for organs at risk [41,42,44–46]. However, the application of deep-learning in tumour segmentation is far from comprehensive and developed. We found only one deep learning study on the auto-segmentation of tumour volume by Men et al., and the accuracy for the segmentation was relatively low [42]. Difficulties with target delineation are mainly determined by wide-ranging differences in tumour volume, location and shape. We expect better accuracy with more training data, but human interference and review will probably remain required for clinical safety [47].

The traditional radiomic features, representing the shape, intensity and texture of ROIs, are well defined by formulas.

These traditional radiomic features are called 'hand crafted' features in many publications to indicate that those features have been defined manually. In radiomics research the associations between the pre-defined image features and clinical outcome is investigated. However, there could be image patterns with a larger association with clinical outcome that are not included in the set of predefined features. Therefore, researchers are also looking for new image features driven by data (clinical outcomes) with Artificial Intelligence techniques like deep learning. Convolutional neural networks (CNNs), the most popular deep learning tool, usually consists of convolutional layers, max pooling layers and fully connected layers which allow deep learning to detect complex image patterns. The outputs of CNNs are features learned adaptively from clinical outcomes. These characteristics potentially make deep learning an effective tool to find predictive and prognostic image features that may enhance the performance of prediction models with traditional radiomics.

A few deep learning studies with specific focus on feature extraction and selection were successful in HNC outcome prediction. Diamant et al. showed that CNNs were capable of improving the area under curve (AUC, 0.86) of the Vallieres model to 0.92 in predicting distant metastasis [25,43]. A hybrid predictive model that combined traditional radiomics and CNN features was proposed by Zhou et al., which showed a better performance than traditional radiomic features and CNN features alone [48]. However, in contrast to the traditional models that take medical indications into account, the automatically obtained deep learning models are difficult to interpret clinically due to the nature of a black box. This limits the use of deep learning in medical image analysis. A lot of research on visualization of deep learning features and models is ongoing to diminish the black box perception [49,50]. We expect that the improvement of interpretability of deep learning features will stimulate the clinical implementation of deep learning models in the future. Another limitation of deep learning is the large risk of overfitting because of the very large numbers of variables. Deep learning can generate a model with a perfect performance on the training set, however, it can fail in an external validation cohort. Overfitting can be reduced by training on a large and representative data set. To define a

Table 2 Radiomic studies in head and neck cancer (HNC) patients.

Year	References	Disease	Image modality	Pts in training	Pts in validation	Endpoint	ROI	No. of features	Software	Validation	Open science
2014	Aerts et al.[3]	HNC	CT	-	231	OS	Manual	440	In-house Matlab	External	Yes (Data)
2015	Leijenaar et al.[6]	OPC	CT	-	542	OS	Manual	440	In-house Matlab	External	Yes (Data)
2015	Parmar et al.[55]	HNC	CT	136	95	OS	Manual	440	In-house Matlab	External	Yes (Data)
2015	Parmar et al.[56]	HNC	CT	101	95	OS	Manual	440	In-house Matlab	External	Yes (Data)
2017	Zhang et al.[57]	NPC	MRI	80	33	PFS	Manual	970	In-house Matlab	Internal	No
2017	Zhang et al.[32]	NPC	MRI	70	40	PFS	Manual	970	In-house Matlab	Internal	No
2017	Vallières et al[25]	HNC	CT/PET	194	106	LRC/DMFS /OS	Manual	1615	In-house Matlab	External	Yes (Data/Code)
2017	Zhai et al.[58]	NPC/HNC	CT	289	298	OS	Manual	139	In-house Matlab	Internal	No
2017	Bogowicz et al. [30]	HNC	CT/PET	121	51	LC	Manual	569	In-house Matlab	External	No
2017	Bogowicz et al. [59]	HNC	CT	93	56	LC	Manual	370	In-house Matlab	External	No

Table 2 Radiomic studies in head and neck cancer (HNC) patients-continued.

Year	References	Disease	Image modality	Pts in training	Pts in validation	Endpoint	ROI	No. of features	Software	Validation	Open science
2017	Ouyang et al.[60]	NPC	MRI	70	30	PFS	Manual	970	In-house Matlab	Internal	No
2017	Zhang et al.[61]	NPC	MRI	80	30	PFS	Manual	970	In-house Matlab	Internal	No
2018	Elhalawani et al.[62]	OPC	CT	465		LRC	Manual	134	IBEX	Internal	No
2018	Feliciani et al.[63]	HNC	PET	90	0	LC	Manual	75	"CGITA"	No	No
2018	Kwan et al.[64]	HPV+ OPC	CT	300	0	DMFS	Manual	Not mention	PyRadiomics /python	Internal	No
2019	Zhai et al.[9]	HNC	CT	240	204	LC/RC/ DMFS/DFS	Manual	120	In-house Matlab	Internal	No
2019	Yuan et al.[65]	HNC	MRI	85	85	OS	Manual	485	In-house Matlab	Internal	No
2019	Cozzi et al.[66]	HNC	CT	70	40	LC/OS	Manual	41	LifeX	Internal	No
2019	Diamant et al.[43]	HNC	CT	194	106	DM	Manual	Deep features	-	External	Yes (Data/ Code)
2019	Leger et al.[39]	HNC	CT	48	30	LRC/OS	Manual	1538	In-house Python	Internal	No

Abbreviations: ROI = region of interest; NPC = nasopharyngeal caner; OPC = oropharyngeal cancer; CT = computed tomography; LC = local control; RC = regional control; DMFS = distant metastasis free survival; DFS = disease free survival; OS = overall survival; MRI = magnetic resonance imaging; PET = positron emission tomography.

representative data set is still a challenge because of large variations that exist between patients. However, deep learning could advance the application of radiomics with a series of challenges in the years to come.

Quality of data and models

In order to assess the quality of current radiomic studies, we performed a literature search in MEDLINE and EMBASE using ((radiomic) OR (radiomics) OR (image biomarker)) AND (head and neck) as keywords. Among 210 unique records, 20 full-text articles (2 studies from our group) reported HNC treatment outcome prediction with models including radiomic features before August 22 2019. Table 2 summarizes the characteristics of these 20 studies. The majority of studies addressed head and neck cancer (13 of 20 studies), 3 studies were specifically on oropharyngeal cancer and 5 studies were on nasopharyngeal cancer.

All radiomic features were extracted from manually contoured structures and more than 100 radiomic features were investigated in 80% of studies. In 15 studies, in-house developed software was used for feature extraction. Open-source package and software such as IBEX and PyRadiomics were used in 5 studies since 2018. In general, all studies provided adequate information on features extraction including image acquisition setting, segmentation and software implementation. The well documented features extraction processes showed large variations between studies. Firstly, different image acquisition parameters and CT scanners of different manufacturers were used in different studies. It is worth mentioning that all studies used retrospective data sets, which means that the scan parameters could also be different within the study. Zhao et al. assessed the stability of radiomic features by using same-day repeated CT scans with different reconstruction algorithms and slice thicknesses [51]. They found that only 19% of the features were repeatable when different settings were used. This indicates that radiomic features may be not reliable when the imaging protocol is not consistent. It is unclear whether the use of different scanners have an even larger influence on the quality and reliability of the extracted radiomic features. Secondly, the manual delineation is susceptible to inter-observer variations. We observed that 9% of radiomic features had inter-observer agreement lower than 0.7. This is in line with the finding of Leijenaar et al.[9,52]. The radiomic features with low inter-observer stability are mainly

intensity and geometric features. Textural features were overall more stable. Thirdly, different in-house developed software tools are used for radiomic feature extraction. The number of investigated features in the 20 studies ranged from 41 to 1,615, the variations between formulas are unknown. The currently available open-source packages can narrow the differences between features extracted from in-house software tools with slightly different definitions. Therefore, the use of validated software tools should be encouraged in the future analysis.

All these variations in feature extraction methods make it difficult to compare different studies and to validate and implement published models. Therefore, it is recommended to develop public image protocol for the standardization of image acquisition in addition to the well documented image protocols [26]. The robustness of radiomic features for inter- and intra- observer variability in contouring needs to be assessed and reported. Furthermore, the "image biomarker standardisation initiative" (IBSI) to standardize the definitions of radiomic features was published in 2016. Using only (open source) software that was developed according to the IBSI definitions is recommended as this will improve the stability and reproducibility of radiomic features between different studies [53].

Only when large datasets with standardized image sets and well defined and stable image features are available can large clinical studies to develop and validate prediction models based on radiomic features be performed. The number of radiomic features are much larger than the size of dataset (Table 2). To reduce the probability of overfitting and multicollinearity, pre-selection was performed in most studies [7–9,33]. Next, multivariable analysis is expected to provide a more holistic model in contrast to the univariable analysis against the endpoints. Generally, the model fitting is optimal in the training set but this may be too optimistic. The TRIPOD statement suggests that a correction of coefficients in the model should be done with internal validation technique [54]. The repeated internal feature selection process can correct for optimism in the models due to overfitting effectively. Finally, an external validation with a dataset from a different institution is necessary for the assessment of the clinical value of the model.

However, two studies performed multivariable analysis without internal validation, 12 (60%) studies did not include external validation due to the lack of datasets from other

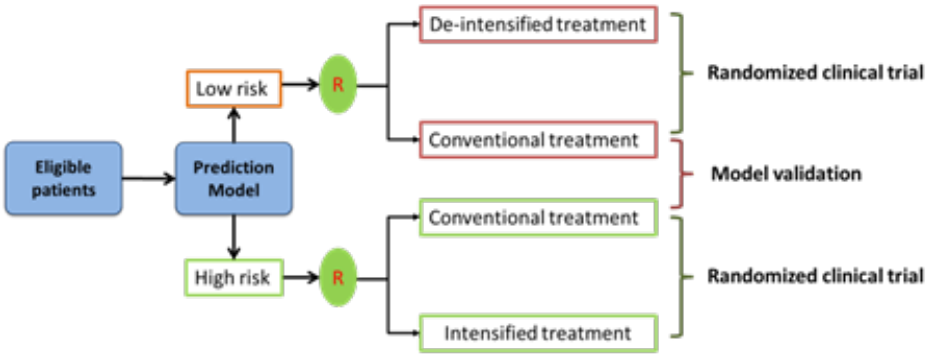

Figure 3. Clinical trial based on model selection. Based on the prediction model, patients can be classified into low- and high-risk groups, patients with different treatment failure risk are selected for different randomized clinical trials. After treatment, the treatment outcomes of the low risk patients and high-risk patients treated with conventional treatment are compared to validate the model. Abbreviation: R = randomization.

institutions. This emphasises the need for data sharing. Shared datasets can markedly improve the options for external validation analysis. Five studies published their data online, one can be obtained by contacting the authors. Two of these studies also made their analysis code publicly available. All of these shared data and codes can accelerate the progress of radiomics.

The lack of standardization of acquisition parameters, inconsistent radiomic extraction methods, and lack of reproducibility of radiomic features are the main limitations in the current radiomic data and models. Researchers are working on overcoming these limitations, which can make radiomics more acceptable in the medical community.

6.2.2. Clinical application:

Many studies have investigated the use of radiomics in prognostic models for head and neck cancer patients (Table 2). The models combining radiomics with classical patient and tumour characteristics allow for more precise risk estimations. Successful external validation of these radiomic models has been demonstrated in several studies [3,6,25,30,43,55,56,59]. This indicates that the models have the potential to be used for personalised treatment-decision making after clinical validation. Using these models, tumours are classified into low- or high-risk groups, treatment strategies can be proposed targeted on those different risk groups. For example, intensified treatment schemes for tumours with a high treatment failure risk can be applied to improve tumour control, while less intensive treatment can be an option for low risk groups to reduce normal tissue side effects.

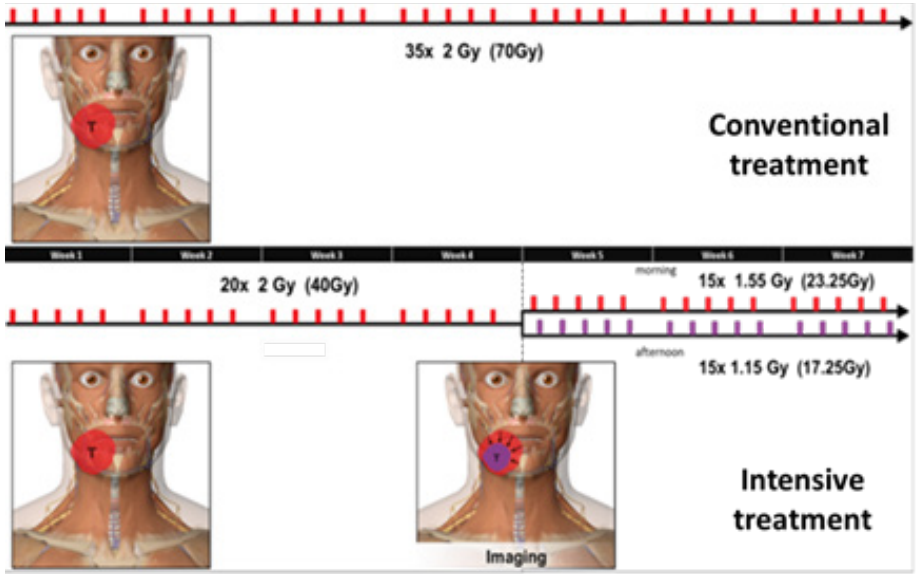

Figure 4. Possible clinical trial based on the local control model. (Courtesy of L.V. van Dijk)

Before the new treatment schemes can be introduced into clinical practice, clinical trials need to be conducted to test the effectiveness of the new treatment methods. In order to increase the radiation dose to the tumour, proton therapy can be suggested in the clinical trial, which can ideally increase the dose to the target and limit the radiation dose to the normal tissue. However, if head and neck cancer patients are randomized without considering their treatment failure risk, the patients that are expected to benefit most from the new intensive treatment strategy may not be selected for the new treatment, and the patients with low treatment failure risk may be over-treated with proton therapy unnecessarily. Therefore, the models based on radiomic features and classical prognostic variables can be used to identify the patients with high treatment failure risk (Figure 3) eligible for the randomized controlled trials (RCTs) of intensified treatment strategies [67]. Similarly, patients with low treatment failure risk identified by the models will be eligible for the RCTs to investigate de-intensified treatment schemes.

Eligible patient selection + Clinical trial

We recommend using the models as eligible criteria for RCTs aiming at introducing new intensified and de-intensified new treatments. The principal is depicted in Figure 3. By calculating the treatment failure risk, patients can be classified into low- and high-risk groups. In this way, patients with a low treatment failure risk can be prevented from

intensive treatment strategies that are only necessary for the high failure risk group. Further randomized controlled trials can help us to find intensified radiation schemes with proton therapy for the high failure risk group and de-intensified radiation scheme for the low failure risk group. Considering the limited availability and high costs of proton treatment, we can demonstrate the benefit of intensified treatment strategies in a smaller number of patients enriched with high risk patients using the models. This strategy is more efficient and has great economic benefits. The comparison between low-risk and high-risk patients treated with conventional treatment can be used to clinically validate the based model in reverse.

We have started a clinical trial based on the local control model described in **Chapter 3** in collaboration with the MD Anderson Cancer Center (MDACC), initiated by van Dijk et al. from UMCG. The local control model was validated and optimised on MDACC data. We estimated the local recurrence risk of patients based on the optimised model. A total of 18 HNC patients with estimated local recurrence risk higher than 60% are enrolled in a Phase I trial to test the safety and feasibility of a novel radiation dose escalation regime. In this starting phase Istudy, the maximum tolerable dose administered to the tumour with proton therapy without severe unacceptable side effects will be determined. If this phase I trial is successful, the resulting next step will be to initiate a multi-center phase II randomized controlled trial within UMCG and MDACC. In Figure 4, the design of a possible future phase II clinical trial is presented. For the conventional treatment group, patients will receive 35 daily fractions of 2 Gy in 7 weeks. In the first 4 weeks of the intensive treatment group, patients will receive conventional daily treatment with 2 Gy. By the end of week 4, mid-treatment imaging will be performed to assess the geometric change of tumour. Based on the decreased tumour target, a hyper-fractionated (twice-daily) radiation plan will be scheduled using 1.55 Gy on the original tumour region plus a boost dose of 1.15 Gy on the decreased tumour target. We will assess the tumour control rate in the prospective follow-up.

The nodal control model in **Chapter 4 and 5** showed promising performance in predicting individual nodal failure risk, and was successfully validated in **Chapter 5**. This model could be used in a clinical trial aiming at improving neck management in HNC. Before such an RCT can start, external validation on larger datasets and a feasibility

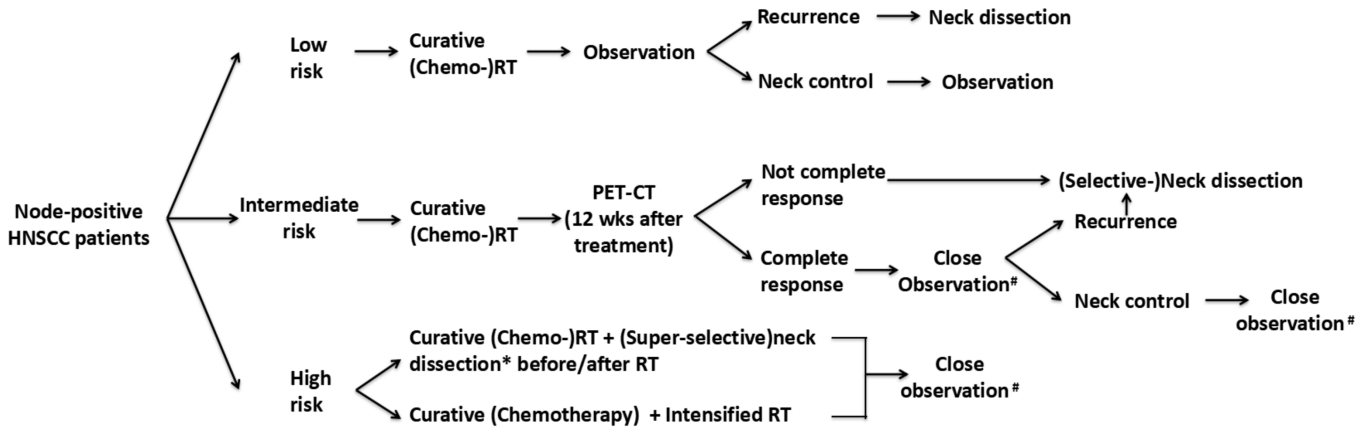

Figure 5 Recommended neck management workflow for node positive head and neck squamous carcinoma patients. #Close observation suggests imaging follow up every 3-6 months in the first 2 years after treatment
*Super-selective neck dissection: to remove only the lymph nodes that are at high risk according to the model
Abbreviations: HNSCC = head and neck squamous cell carcinoma; RT = radiotherapy; PET-CT = positron emission tomography – computed tomography.

trial should be performed. Therefore, we only showed some proof of concept clinical new treatments in Figure 5. Based on the nodal control model, pathological lymph node can be identified as low-, intermediate, and high-risk of failure lymph nodes. As a potential future strategy, lymph nodes with low- and intermediate-risk of failure could be followed with a wait-and-see policy instead of surgical dissection in case of clinical residual disease shortly after radiotherapy, with the surveillance of PET-CT. For the high-risk lymph nodes, an intensified radiation schedule or lymph node targeted dissection before or after (chemo-)radiation could be arranged to avoid complex clinical decisions on re-irradiation or severe post-operative complications. Such a workflow might improve the nodal control rate in the high-risk patients and reduce the number of unnecessary lymph node dissections in the low-risk patients.

In summary, radiomics features provide additional information to predict treatment outcome when combined with classical patient and tumour characteristics. Future improvement can be expected by investigating multiple imaging modalities and using more advanced analytical methodologies like artificial intelligence. In particular, this field requires a renewed focus on standardization, interpretability and data sharing. Ongoing and upcoming clinical trials may bring radiomics into clinical application in personalized medicine.

References

1. Leijenaar RTH. Radiomics: Images are more than meet the eyes. Maastricht Datawyse / Univ Pers Maastricht 2017.
2. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016;278:563–77.
3. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5.
4. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RGPM, Granton P, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6.
5. Pfister DG, Spencer SA, Brizel DM, Burtness BA, Busse PM, Caudell JJ, et al. Head and neck cancers, version 1.2015 featured updates to the NCCN guidelines. JNCCN J Natl Compr Cancer Netw 2015;13:847–56.
6. Leijenaar RTH, Carvalho S, Hoebers FJP, Aerts HJWL, Van Elmpt WJC, Huang SH, et al. External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma. Acta Oncol (Madr) 2015;54:1423–9.
7. Benjamini Y., Hochberg Y. Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing Author (s): Yoav Benjamini and Yosef Hochberg Source : Journal of the Royal Statistical Society . Series B (Methodological), Vol . 57 , No . 1 Published by : J R Stat Soc Ser B 1995;57:289–300.
8. van Der Schaaf A, Xu CJ, Van Luijk P, Van'T Veld AA, Langendijk JA, Schilstra C. Multivariate modeling of complications with data driven variable selection: Guarding against overfitting and effects of data set size. Radiother Oncol 2012;105:115–21.
9. Zhai TT, Langendijk JA, van Dijk L V., Halmos GB, Witjes MJH, Oosting SF, et al. The prognostic value of CT-based image-biomarkers for head and neck cancer patients treated with definitive (chemo-)radiation. Oral Oncol 2019;95:178–86.
10. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. Nature 2013;501:338–45.
11. Junttila MR, De Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. Nature 2013;501:346–54.

12. Pickering CR, Shah K, Ahmed S, Rao A, Frederick MJ, Zhang J, et al. CT imaging correlates of genomic expression for oral cavity squamous cell carcinoma. Am. J. Neuroradiol., vol. 34, 2013, p. 1818–22.
13. Maddika S, Ande SR, Panigrahi S, Paranjothy T, Weglarczyk K, Zuse A, et al. Cell survival, cell death and cell cycle pathways are interconnected: Implications for cancer therapy. Drug Resist Updat 2007;10:13–29.
14. Zhu Y, Mohamed ASR, Lai SY, Yang S, Kanwar A, Wei L, et al. Imaging-Genomic Study of Head and Neck Squamous Cell Carcinoma: Associations Between Radiomic Phenotypes and Genomic Mechanisms via Integration of The Cancer Genome Atlas and The Cancer Imaging Archive. JCO Clin Cancer Informatics 2019:1–9.
15. Chen LL, Nolan ME, Silverstein MJ, Mihm MC, Sober AJ, Tanabe KK, et al. The impact of primary tumor size, lymph node status, and other prognostic factors on the risk of cancer death. Cancer 2009;115:5071–83.
16. Miles KA. Tumour angiogenesis and its relation to contrast enhancement on computed tomography : a review. Eur J Radiol 1999;30:198–205.
17. Muz B, Azab AK. The role of hypoxia in cancer progression , angiogenesis , metastasis , and resistance to therapy. Hypoxia 2015:83–92.
18. Gottgens E, Ostheimer C, Span PN, Bussink J, Hammond EM. HPV , hypoxia and radiation response in head and neck cancer. Br J Radiol 2018;91:20180047.
19. Bredholt G, Mannelqvist M, Stefansson IM, Birkeland E, Bø TH, Øyan AM, et al. Tumor necrosis is an important hallmark of aggressive endometrial cancer and associates with hypoxia, angiogenesis and inflammation responses. Oncotarget 2015;6:39676–91.
20. Gerlinger M, Rowan A, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor Heterogeneity and Branched Evolution revealed by multiregion sequencing. N Engl J Med 2012:883–92.
21. Zhang XC, Xu C, Mitchell RM, Zhang B, Zhao D, Li Y, et al. Tumor Evolution and Intratumor Heterogeneity of an Oropharyngeal Squamous Cell Carcinoma Revealed by Whole-Genome Sequencing. Neoplasia 2013;15:1371-IN7.
22. Balluff B, Hanselmann M, Heeren RMA. Mass Spectrometry Imaging for the Investigation of Intratumor Heterogeneity. vol. 134. 1st ed. Elsevier Inc.; 2017.

doi:10.1016/bs.acr.2016.11.008.

23. Siu L, Snyder A, Stower H, Swanton C, Verhaak RGW, Zenklusen JC, et al. Toward understanding and exploiting tumor heterogeneity. Nat Med 2015;21:846–53.
24. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges J-P, et al. Intratumor Heterogeneity Characterized by Textural Features on Baseline 18F-FDG PET Images Predicts Response to Concomitant Radiochemotherapy in Esophageal Cancer. J Nucl Med 2011;52:369–78.
25. Vallières M, Kay-Rivest E, Perrin LJ, Liem X, Furstoss C, Aerts HJWL, et al. Radiomics strategies for risk assessment of tumour failure in head-and-neck cancer. Sci Rep 2017;7.
26. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, De Jong EEC, Van Timmeren J, et al. Radiomics: The bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749–62.
27. Aerts HJWL. The potential of radiomic-based phenotyping in precisionmedicine a review. JAMA Oncol 2016;2:1636–42.
28. Liu Z, Wang S, Dong D, Wei J, Fang C, Zhou X, et al. The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges. Theranostics 2019;9:1303–22.
29. Ger RB, Zhou S, Chi PCM, Lee HJ, Layman RR, Jones AK, et al. Comprehensive Investigation on Controlling for CT Imaging Variabilities in Radiomics Studies. Sci Rep 2018;8:1–14.
30. Bogowicz M, Riesterer O, Stark LS, Studer G, Unkelbach J, Guckenberger M, et al. Comparison of PET and CT radiomics for prediction of local tumor control in head and neck squamous cell carcinoma. Acta Oncol (Madr) 2017;56:1531–6.
31. Chen WS, Li JJ, Hong L, Xing ZB, Wang F, Li CQ. Comparison of MRI, CT and 18F-FDG PET/CT in the diagnosis of local and metastatic of nasopharyngeal carcinomas: An updated meta analysis of clinical studies. Am J Transl Res 2016;8:4532–47.
32. Zhang B, He X, Ouyang F, Gu D, Dong Y, Zhang L, et al. Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma. Cancer Lett 2017;403:21–7.
33. Zhai TT, van Dijk L V., Huang BT, Lin ZX, Ribeiro CO, Brouwer CL, et al. Improving

the prediction of overall survival for head and neck cancer patients using image biomarkers in combination with clinical parameters. Radiother Oncol 2017;124:256–62.

34. van Dijk L V., Thor M, Steenbakkers RJHM, Apte A, Zhai TT, Borra R, et al. Parotid gland fat related Magnetic Resonance image biomarkers improve prediction of late radiation-induced xerostomia. Radiother Oncol 2018;128:459–66.
35. Wu MH, Chen CN, Chen KY, Ho MC, Tai HC, Wang YH, et al. Quantitative analysis of echogenicity for patients with thyroid nodules. Sci Rep 2016;6.
36. Moon WK, Huang YS, Lo CM, Huang CS, Bae MS, Kim WH, et al. Computer-aided diagnosis for distinguishing between triple-negative breast cancer and fibroadenomas based on ultrasound texture features. Med Phys 2015;42:3024–35.
37. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and Reproducibility of Radiomic Features: A Systematic Review. Int J Radiat Oncol Biol Phys 2018;102:1143–58.
38. Larue RTHM, Defraene G, De Ruysscher D, Lambin P, Van Elmpt W. Quantitative radiomics studies for tissue characterization: A review of technology and methodological procedures. Br J Radiol 2017;90.
39. Leger S, Zwanenburg A, Pilz K, Zschaeck S, Zöphel K, Kotzerke J, et al. CT imaging during treatment improves radiomic models for patients with locally advanced head and neck cancer. Radiother Oncol 2019;130:10–7.
40. 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 Biol Phys 2019;104:942–52.
41. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks: Med Phys 2017;44:547–57.
42. Men K, Chen X, Zhang Y, Zhang T, Dai J, Yi J, et al. Deep deconvolutional neural network for target segmentation of nasopharyngeal cancer in planning computed tomography images. Front Oncol 2017;7.
43. Diamant A, Chatterjee A, Vallières M, Shenouda G, Seuntjens J. Deep learning in head & neck cancer outcome prediction. Sci Rep 2019;9.
44. Fritscher K, Raudaschl P, Zaffino P, Spadea MF, Sharp GC, Schubert R. Deep neural

networks for fast segmentation of 3D medical images. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9901 LNCS, 2016, p.158–65.

45. Močnik D, Ibragimov B, Xing L, Strojan P, Likar B, Pernuš F, et al. Segmentation of parotid glands from registered CT and MR images. Phys Medica 2018;52:33–41.
46. Ren X, Xiang L, Nie D, Shao Y, Zhang H, Shen D, et al. Interleaved 3D-CNNs for joint segmentation of small-volume structures in head and neck CT images. Med Phys 2018;45:2063–75.
47. Kosmin M, Ledsam J, Romera-Paredes B, Mendes R, Moinuddin S, de Souza D, et al. Rapid advances in auto-segmentation of organs at risk and target volumes in head and neck cancer. Radiother Oncol 2019;135:130–40.
48. Zhou Z, Chen L, Sher D, Zhang Q, Shah J, Pham NL, 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 2019;64.
49. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. Proc. IEEE Int. Conf. Comput. Vis., vol. 2017- Octob, 2017, p. 618–26.
50. Chattopadhay A, Sarkar A, Howlader P, Balasubramanian VN. Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks. Proc. - 2018 IEEE Winter Conf. Appl. Comput. Vision, WACV 2018, vol. 2018- Janua, 2018, p. 839–47. doi:10.1109/WACV.2018.00097.
51. Zhao B, Tan Y, Tsai WY, Qi J, Xie C, Lu L, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. Sci Rep 2016;6.
52. Leijenaar RTH, Carvalho S, Velazquez ER, Van Elmpt WJC, Parmar C, Hoekstra OS, et al. Stability of FDG-PET Radiomics features: An integrated analysis of test-retest and inter-observer variability. Acta Oncol (Madr) 2013;52:1391–7.
53. Zwanenburg A, Leger S, Vallières M, Löck S, Initiative for the IBS. Image biomarker standardisation initiative 2016. doi:10.17195/candat.2016.08.1.
54. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. Eur Radiol 2015;67:1142–51.

55. Parmar C, Leijenaar RTH, Grossmann P, Velazquez ER, Bussink J, Rietveld D, et al. Radiomic feature clusters and Prognostic Signatures specific for Lung and Head &neck cancer. Sci Rep 2015;5:1–10.
56. 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.
57. Zhang B, Tian J, Dong D, Gu D, Dong Y, Zhang L, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. Clin Cancer Res 2017;23:4259–69.
58. Zhai TT, van Dijk L V., Huang BT, Lin ZX, Ribeiro CO, Brouwer CL, et al. Improving the prediction of overall survival for head and neck cancer patients using image biomarkers in combination with clinical parameters. Radiother Oncol 2017;124:256–62.
59. Bogowicz M, Riesterer O, Ikenberg K, Stieb S, Moch H, Studer G, et al. Computed Tomography Radiomics Predicts HPV Status and Local Tumor Control After Definitive Radiochemotherapy in Head and Neck Squamous Cell Carcinoma. Int J Radiat Oncol Biol Phys 2017;99:921–8.
60. Ouyang FS, Guo BL, Zhang B, Dong YH, Zhang L, Mo XK, et al. Exploration and validation of radiomics signature as an independent prognostic biomarker in stage III-IVb nasopharyngeal carcinoma. Oncotarget 2017;8:74869–79.
61. Zhang B, Ouyang F, Gu D, Dong Y, Zhang L, Mo X, et al. Advanced nasopharyngeal carcinoma: pre-treatment prediction of progression based on multi-parametric MRI radiomics. Oncotarget 2017;8.
62. Elhalawani H, Kanwar A, Mohamed ASR, White A, Zafereo J, Wong A, 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.
63. Feliciani G, Fioroni F, Grassi E, Bertolini M, Rosca A, Timon G, et al. Radiomic Profiling of Head and Neck Cancer: 18 F-FDG PET Texture Analysis as Predictor of Patient Survival . Contrast Media Mol Imaging 2018;2018:1–8.
64. Kwan JYY, Su J, Huang SH, Ghoraie LS, Xu W, Chan B, et al. Radiomic Biomarkers to Refine Risk Models for Distant Metastasis in HPV-related Oropharyngeal Carcinoma.

Int J Radiat Oncol Biol Phys 2018;102(4):1107-1116.

65. Yuan Y, Ren J, Shi Y, Tao X. MRI-based radiomic signature as predictive marker for patients with head and neck squamous cell carcinoma. Eur J Radiol 2019;117:193–8.
66. Cozzi L, Franzese C, Fogliata A, Franceschini D, Navarria P, Tomatis S, et al. Predicting survival and local control after radiochemotherapy in locally advanced head and neck cancer by means of computed tomography based radiomics. Strahlentherapie Und Onkol 2019:195(9):805-818.
67. Bentzen SM. Randomized controlled trials in health technology assessment: Overkill or overdue? Radiother Oncol 2008;86:142–7.