



**HAL**  
open science

## **Radiomics : A primer for the radiation oncologist**

J.-E. Bibault, L. Xing, P. Giraud, R. El Ayachy, N. Giraud, P. Decazes, A. Burgun

► **To cite this version:**

J.-E. Bibault, L. Xing, P. Giraud, R. El Ayachy, N. Giraud, et al.. Radiomics : A primer for the radiation oncologist. *Cancer/Radiothérapie*, 2020, 24, pp.403 - 410. 10.1016/j.canrad.2020.01.011 . hal-03492367

**HAL Id: hal-03492367**

**<https://hal.science/hal-03492367v1>**

Submitted on 18 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

## **Radiomics: A primer for the radiation oncologist**

### ***Radiomique : une introduction pour l'oncologue radiothérapeute***

Jean-Emmanuel BIBAULT <sup>a\*,b,c</sup>, Lei XING <sup>d</sup>, Paul GIRAUD <sup>c</sup>, Radouane EL AYACHY <sup>c</sup>, Nicolas GIRAUD <sup>e</sup>, Pierre DECAZES <sup>f,g</sup>, Anita BURGUN <sup>b,c,h</sup>, Philippe GIRAUD <sup>a,b</sup>

<sup>a</sup> Radiation Oncology Department, hôpital européen Georges-Pompidou, Assistance publique-hôpitaux de Paris, 20, rue Leblanc, 75015 Paris, France

<sup>b</sup> Université de Paris, 85, boulevard Saint-Germain, 75006 Paris, France

<sup>c</sup> Inserm, UMR 1138, Team 22: Information Sciences to support Personalized Medicine, 15, rue de l'École-de-Médecine, 75006 Paris, France

<sup>d</sup> Laboratory of Artificial Intelligence in Medicine and Biomedical Physics, Stanford University School of Medicine, 875 Blake Wilbur Drive, Stanford, CA 94305-5847, USA

<sup>e</sup> Radiation Oncology Department, CHU de Bordeaux, hôpital Haut-Lévêque, avenue Magellan, 33600 Pessac, France

<sup>f</sup> Nuclear Medicine Department, centre Henri-Becquerel, 1, rue d'Amiens, 76038 Rouen, France

<sup>g</sup> Quantif, EA 4108, université de Rouen, avenue de l'Université, 76801 Saint-Étienne-du-Rouvray, France

<sup>h</sup> Biomedical Informatics and Public Health Department, hôpital européen Georges-Pompidou, Assistance publique-hôpitaux de Paris, 20, rue Leblanc, 75015 Paris, France

**\*Corresponding author:** Jean-Emmanuel Bibault; E-mail: [jean-emmanuel.bibault@aphp.fr](mailto:jean-emmanuel.bibault@aphp.fr)

### **Keywords**

Radiomics; Machine learning; Deep learning; Modeling; Radiation oncology; Clinical oncology

### **Mots clé**

Radiomique ; apprentissage profond ; modélisation ; oncologie ; radiothérapie ; clinique

## **Abstract**

*Purpose.*- Radiomics are a set of methods used to leverage medical imaging and extract quantitative features that can characterize a patient's phenotype. All modalities can be used with several different software packages. Specific informatics methods can then be used to create meaningful predictive models. In this review, we will explain the major steps of a radiomics analysis pipeline and then present the studies published in the context of radiation therapy.

*Methods.*- A literature review was performed on Medline using the search engine PubMed. The search strategy included the search terms "radiotherapy", "radiation oncology" and "radiomics". The search was conducted in July 2019 and reference lists of selected articles were hand searched for relevance to this review.

*Results.*- A typical radiomics workflow always includes five steps: imaging and segmenting, data curation and preparation, feature extraction, exploration and selection and finally modeling. In radiation oncology, radiomics studies have been published to explore different clinical outcome in lung (n=5), head and neck (n=5), esophageal (n=3), rectal (n=3), pancreatic (n=2) cancer and brain metastases (n=2). The quality of these retrospective studies is heterogeneous and their results have not been translated to the clinic.

*Conclusion.*- Radiomics has a great potential to predict clinical outcome and better personalize treatment. But the field is still young and constantly evolving. Improvement in bias reduction techniques and multicenter studies will hopefully allow more robust and generalizable models.

## **Résumé**

*Objectif de l'étude.*- La radiomique est un ensemble de méthodes utilisées pour exploiter l'imagerie médicale et extraire des caractéristiques quantitatives afin de définir le phénotype d'un patient. Toutes les modalités d'imagerie peuvent être utilisées avec différents logiciels. Des méthodes informatiques spécifiques sont nécessaires afin de pouvoir créer des modèles cliniques. Dans cette revue, nous expliquerons les principales étapes de ce type d'analyse puis nous présenterons les principales études réalisées dans ce domaine en oncologie radiothérapie.

*Méthodes.*- Une revue de la littérature a été réalisée dans Medline à l'aide du moteur de recherche PubMed. La stratégie de recherche a inclus les termes «radiotherapy », « radiation oncology » et « radiomics ». La recherche a été réalisée en juillet 2019 et la liste générée a ensuite été explorée afin d'évaluer la pertinence de chaque article pour cette revue.

*Résultats.*- Cinq étapes essentielles sont dénombrées au cours d'une étude de radiomique : l'acquisition et la délimitation de l'image, la préparation des données, l'extraction des caractéristiques, leur exploration et leur sélection et enfin la modélisation. En oncologie radiothérapie, la radiomique a

été mise en œuvre dans le contexte du cancer du poumon (n=5), des cancers ORL (n=5), de l'œsophage (n=3), du rectum (n=3), du pancréas (n=2) et des métastases cérébrales (n=2). La qualité de ces études rétrospectives est hétérogène et leur modèle n'a jamais été employé en pratique clinique.

*Conclusion.*- La radiomique présente un grand potentiel dans la prédiction de l'efficacité et de la toxicité des traitements dans une optique de médecine personnalisée. Cette discipline est toutefois encore jeune et en constante évolution. Des améliorations dans les méthodes de réduction des biais et la coopération dans des études multicentriques devraient permettre d'espérer pouvoir créer des modèles plus robustes et applicables en pratique clinique.

## **1. Introduction**

Imaging plays a central role in radiation oncology, from diagnosis to treatment planning and monitoring. The field is also evolving towards a new direction: with radiomics, images can be used to create digital biomarkers that can be leveraged to personalize treatments. Radiomics can provide a mathematical quantification of tumor phenotype using high-dimension data generated from segmented medical imaging. A very high number of features can be hand-crafted or automatically extracted from computed tomography (CT), magnetic resonance (MR), positron emission tomography (PET), alone or combined, to be later correlated to tumor characteristics and clinical outcomes. These features must also be combined with other pertinent data, such as medical notes from electronic-health records, pathology, biology or genomics, to produce robust models. These models can then be used to improve the patient's outcome. This field is evolving rapidly and holds many promises in radiation oncology. Many challenges still need to be addressed: the adequate curation of quality data, the external validation of the models on independent cohorts and finally the validation of the approach in a randomized controlled phase 3 trial designed to prove the clinical relevance of radiomics. Many models have been created and published, but these studies often lack standardized evaluation on external cohorts of patients. Their reproducibility and generalization are still unsure, which explains why not a single model has been translated to clinical practice.

In this review, we provide the necessary concepts to the radiation oncologists to understand the process involved in radiomics modeling. As the number of studies in the field is posed to steadily increase, we believe radiation oncologists should be able to distinguish the pitfalls, challenges and opportunities of radiomics when they are applied to their specialty.

## **2. Material and methods**

A literature review was performed in the Medline database using the search engine PubMed. The search strategy included the search terms "radiotherapy", "radiation oncology" and "radiomics".

Reference lists of selected articles were hand searched for relevance to this review. The search was conducted in July 2019, and mainly identified studies published between 1997 and 2019.

Search results were judged for relevance using the title, abstract and full text for inclusion in the analysis. Studies were included when they were published as full articles in English employing a radiomics technique in the field of radiation oncology. A preferred reporting items for systematic reviews and meta-analyses (PRISMA) flowchart of the studies selection process is shown in figure 1.

### **3. Results**

#### ***3.1. The workflow of radiomics***

There are many different methods to create a pipeline of radiomics analysis, but these workflows always consist of five steps (figure 2): imaging and segmenting (1), data curation and preparation (2), feature extraction (3), feature exploration and selection (4) and modeling (5).

##### *3.1.1. Imaging and segmenting*

The first step of radiomics is the acquisition of quality images. The protocol used in CT, MR or PET will have a significant impact on all the following steps of the analysis. Ideally, all the images used in a study should be performed with the same machine and the same protocol. If this is not the case, it should be taken into account and corrected in the modeling step of the analysis. In all imaging modalities, vendor, machine and acquisition protocols are believed to be potential bias [1]. The determination of interscanner and intervender variability of the features should be performed before modeling [2]. Since generalizability requires using multicenter cohorts, phantom studies can be used to assess these uncertainties to identify, correct or exclude the features depending on this variability.

Before any extraction can be performed, a volume of interest needs to be defined. The analysis will focus on the pixels and voxels within that volume. In most cases, this volume is the tumor itself, or in some cases, the immediate surroundings of the tumor, the so-called “shell features”[3]. The variability in segmentation can create bias in the later steps of the analysis and should also be accounted for. Several (at least two) segmentations should be performed. Operators should be careful in excluding any irrelevant nearby structures that would undermine the model, such as bone, air, lymph node (figure 3). The features will be extracted from both segmentation and only the most robust features (the one who do not depend on the operator) will be selected for modeling. Methods to obtain reproducible segmentation involve the evaluation by several experts when a manual segmentation is performed or (semi)-automatic segmentation, using deep learning for instance [4].

##### *3.1.2. Data curation and preparation*

This crucial step can be very time-consuming, but needs to be performed rigorously. Several key elements must be defined before beginning any feature extraction or modeling:

- the imaging modality: CT, MR, PET or a combination;
- the imaging protocol: field of view, slices thickness, injection timing;
- the volume of interest: primary tumor, tumoral shell or one or several metastatic lesions;
- the segmentation protocol: manual, or automatic, the number of segmentation needed, the review process;
- the prediction target (the clinical outcome that will be predicted by the analysis): local control, survival or toxicity.

Any and all of these decisions will impact the relevance of the model being created [5,6]. Beyond the radiomics features, other sources of data can be leveraged to create more complex, heterogeneous models, such as clinical phenotypes from electronic health records [7]. The final dataset will need to be correctly structured before the exploratory analysis can begin.

### *3.1.3. Feature extraction*

Radiomics features are quantitative data extracted from images that can be classified into four categories [8] :

- first order features describe the distribution of voxel intensities;
- shape features are related to the shape of the volume;
- texture features can reveal the intra-tumoral heterogeneity;
- wavelets features calculate the intensity and texture features from wavelet decompositions of the original image.

#### *3.1.3.1. Hand-crafted features*

Historically, radiomics features were extracted from the volume of interest using software freely-available online, such as imaging biomarker explorer (IBEX), LIFEx, 3DSlicer, proprietary solutions, or “in-house” implementations of known extraction algorithms [9–16]. A comprehensive review of radiomics software is available [17]. Variability between these packages has rarely been assessed, but the few studies performed showed significant differences between packages used [18,19]. This is raising concerns regarding the reproducibility of any radiomics models [20–22].

#### *3.1.3.2. Deep learning extractors*

The intrareader and interreader variability that can result from manual segmentation and the variation in imaging and processing techniques for feature extraction greatly impact the models that are created. Deep learning offers several solutions in that setting by automating segmentation, extraction and learning of relevant radiographic features, without the need for human intervention, at any phase of the analysis pipeline. For that reason, deep learning could boost reproducibility, generalizability and accuracy and reduce potential bias of a model [23,4,24].

### 3.1.4. Feature exploration

At this point, the dataset should be divided into training and validation datasets, that will be used to create the model, and a testing dataset. The train/validation/test ratios depend on the method of validation (see below), but most of the time 80% of the data will then be used for the training dataset and 20% for the testing dataset. The validation dataset will be resampled as 20% of the training dataset. It will be used to check for overfitting during the hyperparameter tuning of the model. The test dataset will be kept apart and used to validate the model after the training.

All the following steps are performed on the training dataset alone. Performing exploratory analysis or feature selection with the test dataset will bias the model through a phenomenon known as “test leakage” (the model has already learned something from the test data).

The exploration phase consists in assessing the robustness and relevance of the radiomics features that were previously extracted. The selection often involves controlling the inter-reader variability of the features by comparing the results of different segmentation. Only the best features, i.e. the one that are consistent across different segmentation, are included in the next step.

Features highly correlated with existing, simpler, characteristics (clinical, biological features) do not add value to the model and should be eliminated. In the same manner, radiomics features that are correlated between themselves should not be further explored: only one of them can be included.

### 3.1.5. Modeling

#### 3.1.5.1. Feature selection

The features correlated with the outcome that is being explored should be kept for modeling. A good rule-of-thumb to define the number of features that should be included in a model is a ratio of three to ten patients for each feature included in the modeling should be respected. Several dimensionality reduction techniques are available and the best method depend on the analysis pipeline that will be later applied to create the model [25]. Modeling small populations using thousands of features will only find random, irrelevant, noise within the data and will promote overfitting: the model is very-well adapted to the known data, but will be unable to provide any valid prediction on other unknown, new, data (for example from another hospital).

#### 3.1.5.2. Modeling methodology

There are countless methodologies to create a model with radiomics. From regression to deep learning, each has its own advantages and pitfalls [7]. If there are enough patients included in the cohort, solutions leveraging deep learning should be the preferred method in the years to come. Whichever solution has been selected, the model’s performance should be compared to a baseline, simple method.

Creating a highly complex deep learning model that provides performances similar to simpler statistical tests or machine learning algorithm should be avoided.

### 3.1.5.3. Validation

Depending on the number and availability of different patient cohort, a relevant method of validation should be chosen. If no external cohort is available and the number of patients is low (less than a thousand for example), a  $k$ -fold cross validation should be performed. The dataset is resampled into  $k$  equal sets;  $k-1$  sets are used to train the model and the last remaining set is used to assess the model quality. This is done  $k$  times to test each set and the reported metrics is often the mean of the  $k$  iterations. If the number of patients is high, the preferred method of validation is leave-one-out where a percentage of the dataset is kept apart to create a separate test set and assess the model's quality after training. If an external cohort is available, the model is trained on an internal cohort and validated on the external set to assess generalizability.

The quality of the model is often reported using the receiver operating characteristic (ROC) curve or the area under the ROC curve (AUC), which quantifies the sensitivity and specificity of the model. AUC are typically reported with their confidence interval and are between 0 (worst) and 1 (best) [1].

## 3.2. Radiomics in practice

In this part, we will present radiomics studies that have been published in the field of radiation oncology for each type of cancer (table 1).

### 3.2.1. Lung cancer

One of the first robust studies published in 2014 included 1019 patients and explored the value of radiomics for prognosis in lung cancer [8]. Aerts et al. were able to predict survival in two independent external cohorts of patients and outperformed the gold standard of TNM status for patients with stage III non-small cell lung cancer treated with radiation alone or chemoradiation. Their reported concordance index (C-index, a measure of goodness of fit) was 0.65 ( $p= 2.9 \times 10^{-9}$ ). Radiomics have also been used to model outcome after stereotactic body radiation therapy: in 2016, Huynh et al. successfully predicted distant metastasis in an independent validation dataset using radiomics (C-index=0.67,  $p<0.1$ ), even when none of the conventional and clinical features were prognostic [26]. In 2018, Hosny et al. published a study using a convolutional neural network, a type of deep learning, for lung cancer prognostication using seven independent datasets to stratify patients on their mortality risk in order to better personalize treatments [27]. More recently, Xu et al. published a radiomics study relying on deep learning to predict treatment response from serial imaging. Patients ( $n=179$ ) with stage III non-small cell lung cancer treated with chemoradiation were included and a model was developed to predict 2-year survival. AUC was 0.74 ( $p<0.05$ ) [24]. Deep learning has also been used in the study published by Lou et al. to predict the outcome after stereotactic body radiation therapy and



individualize radiation dose, showing that research teams in the field are moving away from traditional hand-crafted radiomics and using more robust and stable deep learning approaches [23].

### 3.2.2. Head and neck cancer

In head and neck, the study published by Aerts et al. in 2014 showed that radiomics modeling also provided good performances with a C-index of 0.69 ( $p=3.5 \times 10^{-6}$ ). Radiomics have also been used to explore a wide range of clinical outcomes [28].

Yu et al. extracted 1683 features with IBEX from head and neck gross tumor volume with or without lymph nodes in order to predict human papilloma virus (HPV) status. A general linear model was used as it provided better AUC compared with various other models (such as random forest, support vector machine [SVM], decision trees, deep learning). The model performance was tested on two external datasets. AUC was 0,87 on the first and 0,92 on the second dataset. Clinical features, such as grade or TNM stage, were not used [29]. Kann et al. developed a convolutional neural network trained on 2875 lymph nodes delineated from the CT scans of 270 patients to determine whether a lymph node was metastatic or not. On an independent test set, the model demonstrated an AUC of 0,91 (95%CI:0,85-0,97) [30]. Ou et al. created a radiomic signature to estimate overall survival [31]. Data from 120 patients with stage III – IVb (TNM 2010) were included. Oncoradiomics™ was used to extract 544 features, narrowed to 24 statistically significant features by logistic regression, from which a radiomic signature score was generated. The signature was able to predict 5-year survival with an AUC = 0,67 95% CI (0,58 – 0,76). Finally, the MD Anderson Cancer Center head and neck quantitative imaging working group has also published a radiomic signature created on 465 patients with oropharyngeal cancer to assess the recurrence probability after intensity-modulated radiation therapy. Using IBEX, 134 radiomic features were extracted from the primary gross tumor volume and narrowed to two features with decision tree modeling. Local tumor control for patients with a low-risk signature was 94% at 5 years, compared to 62% to 80% for the patients with a high-risk signature [32].

### 3.2.3. Esophageal cancer

For esophageal cancer, radiomics was used to predict pathologic complete response to preoperative chemoradiation. Van Rossum et al. created a model with the pretreatment ( $^{18}\text{F}$ )-fluorodeoxyglucose-PET from 217 patients that provided a C-index of 0.77 [33]. Radiomics have also been used to predict lung toxicity after radiation therapy: Cunliffe et al. used serial computed thoracic tomography and showed a correlation between radiomics-based features with radiation therapy dose and radiation pneumonitis development after chemoradiation for esophageal cancer [34]. AUC was between 0.59 and 0.84. In a more recent study, Amyar et al. showed in a population of 97 patients treated by chemoradiotherapy that a three-dimensional (3D) convolutional neural network, applied on the baseline PET to predict the response to the treatment, outperformed bidimensional (2D) convolutional neural network architectures and more classical radiomics approaches (handcrafted feature extraction

with machine learning classifiers), with maximal AUC =  $0.70\pm 0.02$ ,  $0.67\pm 0.08$  and  $0.62\pm 0.04$ , respectively. Moreover, the addition of a margin around the target lesion seemed to increase the accuracy of the 3D convolutional neural network model (AUC= $0.74\pm 0.02$ ) [35].

#### 3.2.4. Rectal cancer

Radiomics have been used in several studies to predict pathologic complete response after neoadjuvant chemoradiation with MRI or PET/CT with logistic regression or simpler artificial neural network, with correct accuracy (AUC=0.71-0.79 for the MRI model, not provided for the PET/CT model) [35-37]. The most robust study included 222 patients to build a radiomics signature with 30 MRI features [36][35]. The model was created with support vector machine and found an AUC of 0.9756 (95% confidence interval, 0.9185-0.9711) in the validation cohort. However, feature extraction from MRI is even more complex and less reproducible than CT scan, meaning that these results cannot be easily reproduced on another cohort [38][37]. For this reason, extracting features from treatment planning CT scans could be both easier and more reproducible. This approach was used and coupled with a deep learning approach with a 80% accuracy [39].

#### 3.2.5. Pancreatic cancer

Only two studies have explored the role of radiomics for pancreatic cancer treated with radiation therapy. In 2016, Cui et al. included 139 patients and extracted features from prestereotactic body radiation therapy ( $^{18}\text{F}$ )-fluorodeoxyglucose PET [40]. They created a signature with seven features to characterize tumor phenotypes, which was significantly associated with overall survival ( $p=0.002$ , hazard ratio 2.74). More recently, Cozzi et al. curated a cohort of 100 patients treated with stereotactic body irradiation in order to create a signature to predict overall survival (C-index 0.73,  $p=0.01$ ) [41].

#### 3.2.6. Brain metastases

Radiomics have been used in two studies for brain metastases: Cha et al. included 110 patients and directly trained a deep learning network (without feature extraction) to predict response after stereotactic body radiation therapy. AUC was 0.856 (68.2%-100%) in their study [42]. Peng et al. used radiomics to distinguish progression from radionecrosis after stereotactic body radiation therapy for brain metastasis. They included 66 patients with 82 lesions to create a signature with 51 features extracted from MR imaging. AUC was 0.81 [43]. Several other studies had similar results with fewer patients included [44,45]. Overall, the technical quality of the studies leveraging radiomics in brain tumors is not comparable to the studies published in lung or head and neck cancer.

## 4. Discussion

In all the studies detailed in paragraphs 3.2, the signatures or scores have been created using retrospective cohorts. Most of the time, these cohorts were small (around hundred patients) and no

external validation was performed. Radiomics as a field is still very young and exploratory. None of the models mentioned are being used in the clinical routine. The science behind radiomics may still seem immature in several regards: the methods used are very heterogeneous, quality control is often non-existent.

Radiomics has a significant potential to improve medical care, but before it can actually deliver on these promises, we will need to better define the methods used, the quality criteria and the way to validate and then monitor the results of a model. The radiomics quality score has been proposed by Lambin et al. in order to better assess the robustness of a model [1]. We will also need national and international collaborations to share data and models [46]. Specific infrastructures and process, like common ontologies [47], will need to be defined because many technical and ethical challenges remain before this can be achieved.

It is also possible that deep learning will actually replace the “traditional” radiomics pipeline altogether. Segmentation, feature extraction and filtering, and finally the modelling itself, could be better standardized through deep learning by limiting the intervention of humans at each of these steps. The interpretability of the models created with deep learning will need to be carefully examined. This can now be somewhat easily done with the use for example of saliency maps that show the area of an image that explains the results of the model. We will rapidly move away from the five steps of modeling explained in this review and integrate them all into a deep learning algorithm. The question whether this will still be called “radiomics” is relevant.

## **5. Conclusion**

Radiomics is a way to extract features from medical imaging and correlate them to clinical outcomes, such as treatment efficacy or toxicity. It could potentially be used to personalize radiation treatments. But the field is still very young and the quality of the studies published is heterogeneous. Several evolutions will keep allowing significant improvements: the promotion of collaboration, the use of multicenter cohorts, the replacement of hand-crafted techniques by automated deep learning algorithms should let us hope that the models created will be robust enough to be translated into clinical practice.

## **References**

- [1] 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. doi:10.1038/nrclinonc.2017.141.
- [2] Mackin D, Fave X, Zhang L, Fried D, Yang J, Taylor B, et al. Measuring computed tomography scanner variability of radiomics features. *Invest Radiol* 2015;50:757–65. doi:10.1097/RLI.0000000000000180.

- [3] Hao H, Zhou Z, Li S, Maquilan G, Folkert MR, Iyengar P, et al. Shell feature: a new radiomics descriptor for predicting distant failure after radiotherapy in non-small cell lung cancer and cervix cancer. *Phys Med Biol* 2018;63:095007. doi:10.1088/1361-6560/aabb5e.
- [4] Hosny A, Aerts HJ, Mak RH. Handcrafted versus deep learning radiomics for prediction of cancer therapy response. *Lancet Digit Health* 2019;1:e106–7. doi:10.1016/S2589-7500(19)30062-7.
- [5] Polan DF, Brady SL, Kaufman RA. Tissue segmentation of computed tomography images using a Random Forest algorithm: a feasibility study. *Phys Med Biol* 2016;61:6553–69. doi:10.1088/0031-9155/61/17/6553.
- [6] Balagurunathan Y, Gu Y, Wang H, Kumar V, Grove O, Hawkins S, et al. Reproducibility and prognosis of quantitative features extracted from CT images. *Transl Oncol* 2014;7:72–87. doi:10.1593/tlo.13844.
- [7] Bibault J-E, Giraud P, Burgun A. Big Data and machine learning in radiation oncology: State of the art and future prospects. *Cancer Lett* 2016. doi:10.1016/j.canlet.2016.05.033.
- [8] Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Commun* 2014;5:4006. doi:10.1038/ncomms5006.
- [9] IBEX\_MDA. SourceForge n.d. <https://sourceforge.net/projects/ibex-mda/> (accessed July 31, 2019).
- [10] Zhang L, Fried DV, Fave XJ, Hunter LA, Yang J, Court LE. IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics. *Med Phys* 2015;42:1341–53. doi:10.1118/1.4908210.
- [11] Ger RB, Cardenas CE, Anderson BM, Yang J, Mackin DS, Zhang L, et al. Guidelines and experience using imaging biomarker explorer (IBEX) for radiomics. *J Vis Exp* 2018. doi:10.3791/57132.
- [12] Nioche C, Orhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. LIFEx: A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res* 2018;78:4786–9. doi:10.1158/0008-5472.CAN-18-0125.
- [13] LIFEx n.d. <https://www.lifexsoft.org/> (accessed July 31, 2019).

- [14] Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging* 2012;30:1323–41. doi:10.1016/j.mri.2012.05.001.
- [15] 3D Slicer n.d. <https://www.slicer.org/> (accessed July 31, 2019).
- [16] Oncoradiomics - Radiomics research Software - Clinical A.I platform n.d. <https://www.oncoradiomics.com/> (accessed July 31, 2019).
- [17] Court LE, Fave X, Mackin D, Lee J, Yang J, Zhang L. Computational resources for radiomics. *Translational Cancer Research* 2016;5:340-348–348. doi:10.21037/8409.
- [18] Foy JJ, Robinson KR, Li H, Giger ML, Al-Hallaq H, Armato SG. Variation in algorithm implementation across radiomics software. *J Med Imaging (Bellingham)* 2018;5:044505. doi:10.1117/1.JMI.5.4.044505.
- [19] 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. doi:10.1016/j.ijrobp.2018.05.053.
- [20] Vial A, Stirling D, Field M, Ros M, Ritz C, Carolan M, et al. The role of deep learning and radiomic feature extraction in cancer-specific predictive modelling: a review. *Translational Cancer Research* 2018;7:803-816–816. doi:10.21037/21823.
- [21] Parekh VS, Jacobs MA. Deep learning and radiomics in precision medicine. *Expert Rev Precis Med Drug Dev* 2019;4:59–72. doi:10.1080/23808993.2019.1585805.
- [22] Coy H, Hsieh K, Wu W, Nagarajan MB, Young JR, Douek ML, et al. Deep learning and radiomics: the utility of Google TensorFlow™ Inception in classifying clear cell renal cell carcinoma and oncocytoma on multiphasic CT. *Abdom Radiol (NY)* 2019;44:2009–20. doi:10.1007/s00261-019-01929-0.
- [23] Lou B, Doken S, Zhuang T, Wingerter D, Gidwani M, Mistry N, et al. An image-based deep learning framework for individualising radiotherapy dose: a retrospective analysis of outcome prediction. *Lancet Digit Health* 2019;1:e136–47. doi:10.1016/S2589-7500(19)30058-5.
- [24] Xu Y, Hosny A, Zeleznik R, Parmar C, Coroller T, Franco I, et al. Deep learning predicts lung cancer treatment response from serial medical imaging. *Clin Cancer Res* 2019;25:3266–75. doi:10.1158/1078-0432.CCR-18-2495.
- [25] Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep* 2015;5:13087. doi:10.1038/srep13087.

- [26] Huynh E, Coroller TP, Narayan V, Agrawal V, Hou Y, Romano J, et al. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. *Radiother Oncol* 2016;120:258–66. doi:10.1016/j.radonc.2016.05.024.
- [27] Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, et al. Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study. *PLOS Medicine* 2018;15:e1002711. doi:10.1371/journal.pmed.1002711.
- [28] Giraud P, Giraud P, Gasnier A, El Ayachy R, Kreps S, Foy J-P, et al. Radiomics and machine learning for radiotherapy in head and neck cancers. *Front Oncol* 2019;9:174. doi:10.3389/fonc.2019.00174.
- [29] Yu K, Zhang Y, Yu Y, Huang C, Liu R, Li T, et al. Radiomic analysis in prediction of human papilloma virus status. *CtRO* 2017;7:49–54. doi:10.1016/j.ctro.2017.10.001.
- [30] Kann BH, Aneja S, Loganadane GV, Kelly JR, Smith SM, Decker RH, et al. Pretreatment identification of head and neck cancer nodal metastasis and extranodal extension using deep learning neural networks. *Sci Rep* 2018;8:14036. doi:10.1038/s41598-018-32441-y.
- [31] Ou D, Blanchard P, Rosellini S, Levy A, Nguyen F, Leijenaar RTH, et al. Predictive and prognostic value of CT based radiomics signature in locally advanced head and neck cancers patients treated with concurrent chemoradiotherapy or bioradiotherapy and its added value to Human Papillomavirus status. *Oral Oncol* 2017;71:150–5. doi:10.1016/j.oraloncology.2017.06.015.
- [32] M. D. Anderson Cancer Center Head and Neck Quantitative Imaging Working Group. Investigation of radiomic signatures for local recurrence using primary tumor texture analysis in oropharyngeal head and neck cancer patients. *Scientific Reports* 2018;8:1524. doi:10.1038/s41598-017-14687-0.
- [33] van Rossum PSN, Fried DV, Zhang L, Hofstetter WL, van Vulpen M, Meijer GJ, et al. The incremental value of subjective and quantitative assessment of 18F-FDG PET for the prediction of pathologic complete response to preoperative chemoradiotherapy in esophageal cancer. *J Nucl Med* 2016;57:691–700. doi:10.2967/jnumed.115.163766.
- [34] Cunliffe A, Armato SG, Castillo R, Pham N, Guerrero T, Al-Hallaq HA. Lung texture in serial thoracic computed tomography scans: correlation of radiomics-based features with radiation therapy dose and radiation pneumonitis development. *Int J Radiat Oncol Biol Phys* 2015;91:1048–56. doi:10.1016/j.ijrobp.2014.11.030.

- [35] Amyar A, Ruan S, Gardin I, Herault R, Clement C, Decazes P, et al. Radiomics-net: Convolutional Neural Networks on FDG PET Images for predicting cancer treatment response. *J Nucl Med* 2018;59:324–324.
- [36] Liu Z, Zhang X-Y, Shi Y-J, Wang L, Zhu H-T, Tang Z, et al. Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Clin Cancer Res* 2017. doi:10.1158/1078-0432.CCR-17-1038.
- [37] Lovinfosse P, Polus M, Van Daele D, Martinive P, Daenen F, Hatt M, et al. FDG PET/CT radiomics for predicting the outcome of locally advanced rectal cancer. *Eur J Nucl Med Mol Imaging* 2017. doi:10.1007/s00259-017-3855-5.
- [38] Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012;30:1234–48. doi:10.1016/j.mri.2012.06.010.
- [39] Bibault J-E, Giraud P, Durdux C, Taieb J, Berger A, Coriat R, et al. Deep learning and radiomics predict complete response after neo-adjuvant chemoradiation for locally advanced rectal cancer. *Sci Rep* 2018;8:12611. doi:10.1038/s41598-018-30657-6.
- [40] Cui Y, Song J, Pollom E, Alagappan M, Shirato H, Chang DT, et al. Quantitative analysis of (<sup>18</sup>F)-fluorodeoxyglucose positron emission tomography identifies novel prognostic imaging biomarkers in locally advanced pancreatic cancer patients treated with stereotactic body radiation therapy. *Int J Radiat Oncol Biol Phys* 2016;96:102–9. doi:10.1016/j.ijrobp.2016.04.034.
- [41] Cozzi L, Comito T, Fogliata A, Franzese C, Franceschini D, Bonifacio C, et al. Computed tomography based radiomic signature as predictive of survival and local control after stereotactic body radiation therapy in pancreatic carcinoma. *PLoS ONE* 2019;14:e0210758. doi:10.1371/journal.pone.0210758.
- [42] Cha YJ, Jang WI, Kim M-S, Yoo HJ, Paik EK, Jeong HK, et al. Prediction of response to stereotactic radiosurgery for brain metastases using convolutional neural networks. *Anticancer Res* 2018;38:5437–45. doi:10.21873/anticancer.12875.
- [43] Peng L, Parekh V, Huang P, Lin DD, Sheikh K, Baker B, et al. Distinguishing true progression from radionecrosis after stereotactic radiation therapy for brain metastases with machine learning and radiomics. *Int J Radiat Oncol Biol Phys* 2018;102:1236–43. doi:10.1016/j.ijrobp.2018.05.041.
- [44] Lohmann P, Stoffels G, Ceccon G, Rapp M, Sabel M, Filss CP, et al. Radiation injury vs. recurrent brain metastasis: combining textural feature radiomics analysis and standard parameters may

increase (<sup>18</sup>F)-FET PET accuracy without dynamic scans. *Eur Radiol* 2017;27:2916–27.  
doi:10.1007/s00330-016-4638-2.

[45] Lopez CJ, Nagornaya N, Parra NA, Kwon D, Ishkanian F, Markoe AM, et al. Association of radiomics and metabolic tumor volumes in radiation treatment of glioblastoma multiforme. *Int J Radiat Oncol Biol Phys* 2017;97:586–95. doi:10.1016/j.ijrobp.2016.11.011.

[46] Foy J-P, Durdux C, Giraud P, Bibault J-E. RE: The rise of radiomics and implications for oncologic management. *J Natl Cancer Inst* 2018;110:1275–6. doi:10.1093/jnci/djy037.

[47] Bibault J-E, Zapletal E, Rance B, Giraud P, Burgun A. Labeling for big data in radiation oncology: the radiation oncology structures ontology. *PLoS ONE* 2018;13:e0191263.  
doi:10.1371/journal.pone.0191263.



**Figure legends**

**Figure 1.** Review of studies on radiomics in the context of radiation therapy: Preferred reporting items for systematic reviews and meta-analyses (PRISMA) flowchart for studies selection.

**Figure 2.** The radiomics workflow.

**Figure 3.** Radiomics for radiation therapy: An example of multimodal segmentation for locally-advanced rectal cancer showing the importance of excluding air, stool or any other irrelevant structure from the volume of interest.

**Identification**

Records identified through  
database searching  
(n = 175)

Additional records identified  
through other sources  
(n = 4)

**Screening**

Records after duplicates removed  
(n = 152)

Records screened  
(n = 152)

Records excluded  
(n = 85)

**Eligibility**

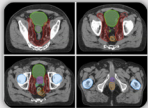
Full-text articles assessed  
for eligibility  
(n = 67)

Full-text articles excluded  
(n = 20)

**Included**

Studies included in  
qualitative synthesis  
(n = 47)

## Imaging and segmenting



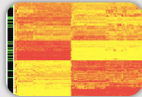
## Data curation and preparation

	A	B	C	D	E
1					
2					
3		Jan	Feb	Mar	Apr
4	Person 1	400	700	250	
5	Person 2	2	820	280	
6	Person 3	700	880	880	
7	Person 4	200	140	310	
8	Person 5	440	500	310	
9	Person 6	400	640	340	
10	Person 7	600	710	310	
11	Person 8	120	480	80	
12	Person 9	380	340	300	
13	Total	4000	5380	3880	

## Feature extraction

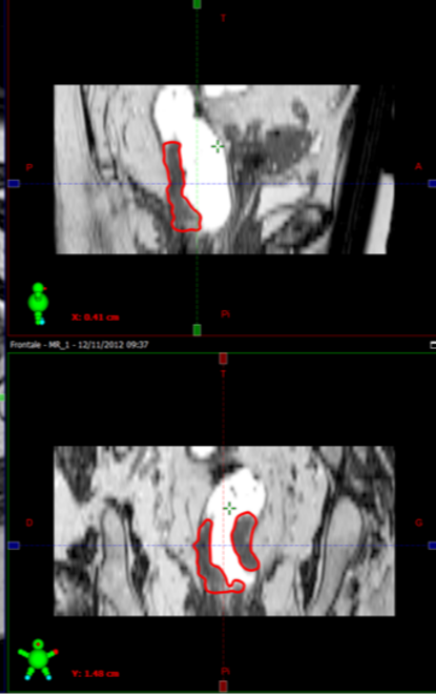
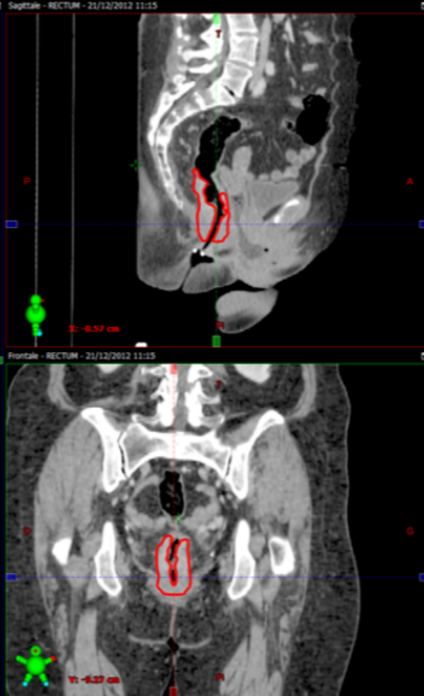


## Feature exploration



## Modeling





**Table 1.** Studies using radiomics to model clinical outcome in radiation oncology.

Cancer	Imaging technique	Number of patients	Features extracted or deep learning methods	Results	Reference
Lung	CT	1,019	Intensity, shape, texture and wavelet	Overall survival after radiotherapy: C-index=0.65, $p= 2.9 \times 10^{-9}$	8
	CT	113	Volume, diameter and LoG 3D RLGL	Overall survival after stereotactic body radiation therapy: C-index = 0.67, $p$ -value<0.1	26
	CT	1,194	deep learning: 3D convolutional neural network	2-year overall survival after chemoradiation: AUC = 0.70 (95% CI: 0.63–0.78), $p < 0.001$	27
	CT	179	deep learning: convolutional neural network	2-year overall survival after stereotactic body radiation therapy: AUC = 0.74 ( $p < 0.05$ )	24
	CT	1,275	deep learning: convolutional neural network	Treatment failure after stereotactic body radiation therapy: C-index = 0.77 (95% CI: 0.69–0.9)	23
Head and neck	CT	1,019	Intensity, shape, texture and wavelet	Overall survival after radiotherapy: C-index=0.69, $p=3.5 \times 10^{-6}$	8
	CT	315	Shape, texture, and grayscale intensities	Assess HPV status: AUC=0.87-0.92	29
	CT	270	deep learning: convolutional neural network	Lymph node metastatic status: AUC=0.91 (95%CI: 0.85-0.97)	30
	CT	120	Oncoradiomics™	Predict 5-year survival: AUC=0.67 (95% CI: 0.58 – 0.76)	31
	CT	465	134 features extracted with IBEX	Signature to predict tumor control: low-risk=94% at 5 years, high-risk=62% to 80%	32
Esophageal	( <sup>18</sup> F)-FDG-PET	217	Total lesion glycolysis, texture and geometry	Complete pathologic response after chemoradiation: C-index=0.77	33
	CT	106	Intensity and texture	Predict radiation pneumonitis: AUC=0.59-0.84	34
	( <sup>18</sup> F)-FDG-PET	97	deep learning: convolutional neural network	Complete response after chemoradiation: AUC=0.74	35
Rectal	MRI	222	30 selected features	Predict pathologic response after neoadjuvant chemoradiation: AUC=0.9756 (95%CI: 0.9185-0.9711)	35
	( <sup>18</sup> F)-FDG-PET	86	Intensity, texture and total lesion glycolysis	Signature associated with disease-free survival	36

	CT	95	Intensity and texture	Predict pathologic response after neoadjuvant chemoradiation: accuracy=80%	38
<b>Pancreas</b>	( <sup>18</sup> F)-FDG-PET	139	Shape, intensity and texture	Signature significantly associated with overall survival (p=0.002, hazard ratio: 2.74) after stereotactic body radiation therapy	39
	CT	100	Texture	Predict overall survival after stereotactic body radiation therapy (C-index = 0.73, p=0.01)	40
<b>Brain metastases</b>	MRI	110	deep learning	Predict response after stereotactic body radiation therapy: AUC=0.856 (68.2%-100%)	41
	MRI	66	Intensity and texture	Distinguish true progression from radionecrosis after stereotactic body radiation therapy: AUC=0.81	42

3D: three-dimensional; 95 % CI: 95 % confidence interval; AUC: area under the curve; C-index: concordance index; CT: computed tomography; FDG: fluorodeoxyglucose; HPV: human papilloma virus; IBEX: imaging biomarker explorer; LoG: laplacian of gaussian; MRI: magnetic resonance imaging; PET: positron emission tomography; RLGL: run low gray level.