REVIEW ARTICLE

Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures

[1]RUBEN T H M LARUE, MSc, [2]GILLES DEFRAENE, MSc, [1]DIRK DE RUYSSCHER, MD, PhD, [1]PHILIPPE LAMBIN, MD, PhD and [1]WOUTER VAN ELMPT, PhD

[1]Department of Radiation Oncology (MAASTRO), GROW—School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, Netherlands
[2]Department of Oncology, Experimental Radiation Oncology, University of Leuven, Leuven, Belgium

Address correspondence to: Mr Ruben T H M Larue
E-mail: *ruben.larue@maastro.nl*

The authors Ruben THM Larue and Gilles Defraene contributed equally to the study.

ABSTRACT

Quantitative analysis of tumour characteristics based on medical imaging is an emerging field of research. In recent years, quantitative imaging features derived from CT, positron emission tomography and MR scans were shown to be of added value in the prediction of outcome parameters in oncology, in what is called the radiomics field. However, results might be difficult to compare owing to a lack of standardized methodologies to conduct quantitative image analyses. In this review, we aim to present an overview of the current challenges, technical routines and protocols that are involved in quantitative imaging studies. The first issue that should be overcome is the dependency of several features on the scan acquisition and image reconstruction parameters. Adopting consistent methods in the subsequent target segmentation step is evenly crucial. To further establish robust quantitative image analyses, standardization or at least calibration of imaging features based on different feature extraction settings is required, especially for texture- and filter-based features. Several open-source and commercial software packages to perform feature extraction are currently available, all with slightly different functionalities, which makes benchmarking quite challenging. The number of imaging features calculated is typically larger than the number of patients studied, which emphasizes the importance of proper feature selection and prediction model-building routines to prevent overfitting. Even though many of these challenges still need to be addressed before quantitative imaging can be brought into daily clinical practice, radiomics is expected to be a critical component for the integration of image-derived information to personalize treatment in the future.

INTRODUCTION

The use of quantitative imaging has been an attractive field of research to overcome the subjectivity of visual interpretation. However, in all imaging divisions from radiology to nuclear medicine, the amount of quantification is still limited and the majority of clinical decision-making is based on visual assessment. A common quantification is performed by the response evaluation criteria in solid tumours[1] that is based on the measurement of tumour size and frequently used for response assessment in oncology, whereas PET Response Criteria in Solid Tumours[2] is making its introduction in the nuclear medicine arena to allow simple quantification of maximum uptake of a tracer [*e.g.* maximum standardized uptake value (SUV) or SUV peak].

Besides these simple quantification methods, diagnosis is also complemented visually by the appearance of lesions having different properties or patterns that are used to differentiate between benign and malignant lesions. These appearances (*i.e.* imaging features) are typically described visually and the radiologist interprets and selects suspected lesions for future clinical investigations (*e.g.* biopsy). For many years, researchers have investigated computer-aided diagnosis techniques to automatize the workflow and improve accuracy.[3–5]

Nowadays, there is renewed interest in the combination of both quantification and visual assessment to provide a comprehensive quantification of imaging data sets. Instead of reporting only a single quantitative measure or a visual subjective report, image-processing techniques are available to describe many different properties that could be quantified from imaging such as the shape and size of tumours and intensity-based and textural properties with

or without additional filtering in a quantitative way. To overcome the wealth of parameters and information, these derived values are combined with statistical modelling techniques to predict a certain clinical end point (*e.g.* survival, local relapse); this field of research is now commonly called radiomics.[6]

The radiomics workflow starts with image acquisition. After image reconstruction, a region of interest is selected that defines the volume for feature extraction. Calculation of these features is performed by image-processing software sometimes including different pre- and post-processing steps. Next, a statistical model is built that allows the selection of features that are able to predict the outcome parameter (*e.g.* survival). Finally, a validation of the model needs to be performed, preferably external validation. An example of the workflow is shown in Figure 1. With this review, we aim to give an overview of the various technical routines and protocols that are involved in quantitative imaging studies such as radiomics. The influence and technical aspects of the above steps will be described in detail in the next sections.

IMAGE ACQUISITION
Medical images acquired for standard clinical diagnostics, (radiotherapy) treatment planning and follow-up purposes are a source of information for radiomics analyses. In the field of oncology, the most widely used modalities include ultrasound, CT, positron emission tomography (PET) and MRI. Many radiomics studies are relying on retrospective data sets, in which individual image acquisition parameters can be different. These different settings can have an influence on the quality and reliability of the extracted radiomic features, as will be discussed below in detail for the commonly used imaging modalities.
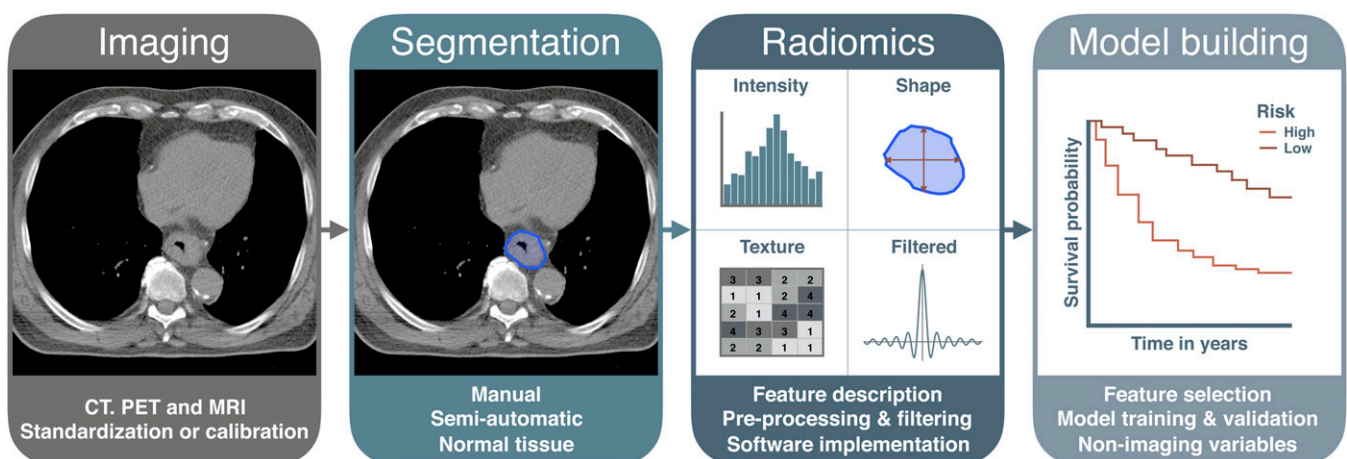
CT
Several groups investigated the repeatability and robustness of radiomic features in CT scans, showing that the features can have a high test–retest stability[7] with an acceptable dynamic range.[8] Mackin et al[9] scanned a phantom with different parameters on CT scanners of four different manufacturers. They found that variability in textural features calculated on CT scans from different scanners can be in the same order of magnitude as the variability observed in CT scans from patients with non-small-cell lung cancer (NSCLC). Zhao et al[10] used the publically available Reference Image Database to Evaluate Therapy Response data set,[11] consisting of 31 patients with NSCLC with same-day repeated CT scans, to assess the impact of slice thickness and reconstruction algorithm on the stability of 89 radiomic features. They concluded that repeatability of features derived from scans with the same imaging settings was good; however, only 19% of the features were repeatable when different settings were used. In conebeam CT scans, Fave et al[12] found that radiomic features may be reliable as long as the imaging protocol is consistent and relative differences are used. In addition to all imaging parameters that should be taken into account before performing radiomics analyses, the influence of respiratory motion on radiomic features should not be underestimated. It was shown that up to almost 75% of the CT radiomic features can be susceptible to respiration, making breath-hold or four-dimensional CT acquisition necessary for moving lesions.[13]

Positron emission tomography
PET scans were used to show a good test–retest stability in up to 71% of the radiomic features in a cohort of patients with NSCLC.[14] In oesophageal cancer, heterogeneity parameters such as entropy, homogeneity, dissimilarity (local characterization) and variability in the size and intensity of homogeneous tumour areas (regional characterization) also had a good reproducibility.[15] However, various studies found that PET radiomic features can be susceptible to reconstruction parameters. For instance, Galavis et al[16] showed that in a cohort of 20 patients with different types of solid tumours, 40 (80%) of the 50 features tested presented large variations (range >30%) when the number of iterations, grid size, reconstruction algorithm and/or post-reconstruction filter was changed. This was also confirmed by van Velden et al[17] and Yan et al,[18] who tested the impact of the same reconstruction parameters on feature robustness. Respiratory motion causing blurring is another matter of concern. Yip et al[19] demonstrated that textural features can vary up to 19% when comparing their values derived from

Figure 1. An overview of the radiomics workflow and corresponding topics addressed in this review. PET, positron emission tomography.

three-dimensional (3D)-PET with those from four-dimensional PET scans. Also, the studies by Oliver et al[13] and Grootjans et al[20] showed that respiratory motion has a significant effect on the quantification of tumour heterogeneity with PET. In addition to this, it is shown that the specific way of SUV discretization has a crucial effect on the resulting textural features, and obviously their interpretation, in both oesophageal[21] and lung carcinomas.[22] This effect turned out to be comparable in both carbon-11 choline and fluorine-18 fludeoxyglucose (FDG)-PET imaging.[23] Similar to the current European Association of Nuclear Medicine/EANM Research Ltd harmonization protocol for FDG-PET imaging,[24] a thorough comparison of imaging and reconstruction parameters is warranted before two independent data sets can be compared with each other in a radiomics analysis.

MRI

Radiomics on MRI has not been investigated as extensively as on CT and PET scans. Its potential value is shown by, for instance, Wibmer et al[25] and Gnep et al,[26] both showing that textural features of prostate MRI may differentiate non-cancerous and cancerous prostate tissues and may correlate with biochemical recurrence and Gleason score. However, to our knowledge, to date, no studies investigating the repeatability and stability of radiomic features in MRI have been published. Given the fact that geometrical distortions are quite common in MRI,[27] further research investigating its effect on radiomic feature extraction is needed.

Ultrasound

Quantitative features retrieved from ultrasound images have been mainly shown to be useful to discriminate among normal, malignant and benign tissues.[28] Andrėkutė et al[29] evaluated whether acoustical, textural and shape features were able to differentiate malignant melanoma from benign melanocytic tumours. They identified a combination of seven relevant features, yielding an accuracy of 82.4%. Similar accuracies were observed when using quantitative (textural) features to identify malignant thyroid nodules[30,31] or breast tumours.[32,33] Although a study by Nadeau et al[34] revealed a high interobserver variability in quantitative ultrasound features of the Achilles tendons, to date no integrated analysis testing the repeatability and stability of ultrasound radiomic features for applications in oncology has been published.

Standardization or calibration

Currently, many radiomics studies in different cancer types have used data acquired in a single institute, sometimes combined with an internal validation step.[26,35–40] These studies showed the great potential for radiomics. However, the lack of standardization/harmonization or at least a correlation between radiomic features acquired in different settings (*e.g.* scanner type, hospital, radiomics software) makes it difficult to directly compare different studies and extracted feature values. This makes blindly interchanging data sets for validation without taking the different acquisition settings into account difficult. Specific radiomic correction and calibration algorithms could potentially solve this issue and are currently being investigated.

For future prospective studies, it could therefore be beneficial to adopt acquisition and reconstruction standards, as proposed by, for instance, the Quantitative Imaging Biomarker Alliance, Quantitative Imaging Network, American Association of Physicists in Medicine and European Association of Nuclear Medicine.[24,41,42] Nonetheless, standardization can be challenging with the introduction of new, state-of-the-art imaging equipment in different institutes and cannot be applied to the enormous amount of retrospective data available.

TARGET VOLUME DEFINITION

Manual segmentation

Accurate segmentation is an important step of the radiomics workflow, as radiomic features are derived from segmented volumes of interest. Manual delineation of the gross tumour volume is a standard clinical routine in the treatment-planning process for patients receiving radiotherapy, but for other interventions, this is not frequently performed. Manual delineation is a straightforward solution, but can also be very time consuming and is susceptible to interobserver variability. For instance, interobserver delineation in cervical cancer can lead to significant differences and are reported to differ up to 4 cm.[43] These differences can influence radiomic features extracted from the delineated volumes. Van Velden et al[17] investigated the influence of reconstruction and delineation on the repeatability of radiomic features in patients with NSCLC and concluded that 24% of the features were susceptible to the delineation method. Also, Leijenaar et al[14] found in a cohort of 23 patients with NSCLC that radiomic features are susceptible to differences in delineation; however, they observed that 91% of the PET features still had a high interobserver stability.

(Semi-)automatic segmentation

Automatic or semi-automatic segmentation methods are currently investigated extensively to minimize manual input and increase consistency in delineating the regions of interest. In a cohort of patients with hepatocellular cancer, Echegaray et al[44] compared manual delineations with core samples that were obtained by automatically tracing the maximal circle inscribed in the outlines. They showed that the same set of stable features can be retrieved from the core sample, providing as much information as a detailed segmentation. Balagurunathan et al[45] showed that most of the radiomic features have a high reproducibility using an automated, seed-based image analysis program with segmentation performed by a single reader. Parmar et al[46] compared manual slice-by-slice delineations of five physicians with semi-automatic segmentation using the 3D Slicer platform. They showed that radiomic features extracted from 3D Slicer volumes had a significantly higher reproducibility and were more robust than those extracted from manual segmentation. Other groups also reported that semi-automatic segmentation methods have a good correlation with pathology.[47,48]

Depending on the application, more advanced methods such as the fuzzy c-means,[49] fuzzy hidden Markov chains[50] or fuzzy locally adaptive Bayesian (FLAB)[51] segmentation algorithms might be preferred over the discussed methods. Especially for smaller lesions, the FLAB algorithm showed a superior

performance when compared with fuzzy c-means or fuzzy hidden Markov chains.[51,52] Two other studies demonstrated that the FLAB algorithm had a good correlation with pathology[53] and that measures derived from FLAB-segmented volumes had a significantly higher predictive value in oesophageal cancer than the same measures derived from volumes segmented with a fixed threshold segmentation.[54]

Although semi-automatic segmentation methods will outperform manual segmentations in terms of repeatability, radiomic features will still be depending on the segmentation method used, as shown in Figure 2. Using fixed thresholds or the FLAB algorithm results in volumes that can differ up to 51%. This obviously influences volume-related features (*e.g.* energy), but also textural features such as grey-level co-occurrence matrix dissimilarity.

Normal tissue segmentation

The majority of the radiomics studies published to date mainly focused on analyses on malignant lesions for use in prognostic or predictive models. However, radiomics analysis of imaging data might also have other applications. For example, it was shown that a texture-based measure of pulmonary CT scans had a superior performance in discriminating patients with and without Chronic Obstructive Pulmonary Disease when compared with the commonly used densitometric measures.[55] Using screening CTs of the National Lung Screening Trial,[56] Hawkins et al[57] developed a radiomics model consisting of 23 stable features to predict nodules that will become cancerous at 1 and 2 years with accuracies of 80% [area under the curve (AUC 0.83)] and 79% (AUC

0.75), respectively. Also, in patients with early-stage NSCLC who received stereotactic body radiation therapy, a combination of high-risk features on post-stereotactic body radiation therapy CT scans was identified to accurately distinguish between local recurrences and fibrosis.[58] These features were also validated in an independent patient cohort.[59] Similarly, Cunliffe et al[60] analyzed the relationship between radiation dose and development of radiation pneumonitis, with changes in CT radiomic features before and after chemoradiation of patients with oesophageal cancer. They observed a significant change in radiomic features with increasing dose for all features, while for a selection of features this change was significantly related to radiation pneumonitis development. Although not all correlations discussed above are confirmed in independent data sets yet, the first results demonstrate the potential ability of quantitative imaging features to assess the risk for the development of cancer in healthy subjects, local recurrences after initial treatment or radiation-induced side effects in normal tissues.
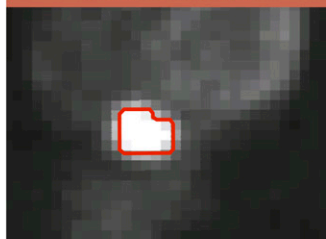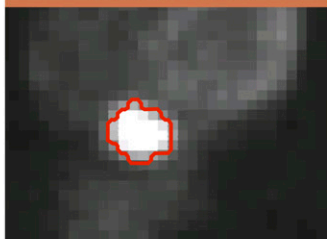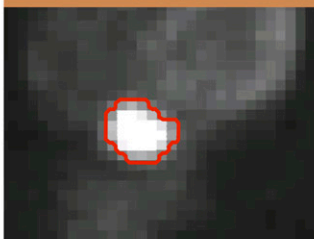
FEATURE EXTRACTION
Radiomic feature description

Radiomic features are derived from the information contained in the voxels of the segmented structure. The features can be grouped into different categories and some typical examples are discussed below. For the full mathematical description, we refer to the respective publications.

First-order statistics are derived from the histogram of voxel intensities (*i.e.* Hounsfield units for CT imaging or SUV values for PET imaging). These histogram characteristics reflect the

Figure 2. Positron emission tomography (PET) radiomic features[14] and their dependency on the delineation method in oesophageal cancer: the 50% maximum standardized uptake value (SUV_{max}) delineation is used as reference. Fuzzy locally adaptive Bayesian (FLAB) delineation was implemented as described previously.[51]

PET feature	50% SUVmax	40% SUVmax		FLAB	
	Value	Value	Change	Value	Change
Volume	26.59	34.61	+30%	40.36	+51%
Entropy	3.64	3.89	+7%	3.98	+10%
Total energy	$4.72 \cdot 10^4$	$5.28 \cdot 10^4$	+12%	$5.57 \cdot 10^4$	+18%
Homogeneity	0.29	0.25	-12%	0.24	-17%
Dissimilarity	2.97	3.60	+21%	3.95	+33%
Zone percentage	0.35	0.39	+9%	0.39	+10%
Size-zone variability	0.36	0.38	+6%	0.38	+4%
Intensity variability	0.09	0.08	-15%	0.07	-18%

mean value, dispersion (standard deviation, mean absolute deviation), central moments (skewness and kurtosis describing asymmetry and sharpness, respectively) and randomness (entropy, uniformity).[61] The area under the cumulative SUV histogram has been used as a heterogeneity measure in FDG-PET scans.[62]

Texture or greyscale variation features, widely used in pattern recognition, refer to higher order statistical measures and summarize the local spatial arrangement of intensities. The textural features are based on different parent matrices capturing this spatial intensity distribution. The grey-level co-occurrence matrix,[63] counting voxel pairs with certain grey values at a predefined direction and distance from each other, generates the features homogeneity, contrast and sum variance. The neighbourhood grey-tone difference matrix[64] is based on the differences between each voxel and the neighbouring voxels, resulting in features said to resemble the human perception of the image: coarseness, complexity and texture strength. Similarly, the neighbouring grey-level dependence matrix[65] considers all neighbouring voxels eliminating angular dependency. Small number emphasis (image fineness) and large number emphasis (image smoothness) are computed from this matrix. Grey-level run length[66] features such as short and long run emphases focus on collinear voxels with the same grey level value. Finally, grey-level size zone-based features[67] target groups of connected pixels with the same grey value.

Wavelet decomposition of the original image has been employed to extract intensity and textural features from different frequency bands[61] and to obtain fused texture characteristics from two imaging modalities.[68] All intensity-based and textural features were mostly calculated in a volumetric way (*i.e.* all voxels within the region of interest were taken into account), with some studies considering only one image slice in the calculation (mostly the slice with the largest cross-sectional area).[69] For the volumetric features, slice thickness clearly impacts on feature extraction results. The nature of matrix-based textural features makes them highly sensitive to the present voxel size anisotropy. Some studies therefore performed a resampling to an isotropic voxel size as a pre-processing step before feature extraction.[68] Others compared the volumetric and in-slice approaches and concluded that the largest cross-sectional area analysis was an acceptable substitute for volumetric analysis for many features. However, in order to avoid potential erroneous conclusions by overlooking parts of the tumour this way, the use of volumetric features incorporating information of the whole tumour was recommended.[70,71]

Shape-based features describe the 3D geometrical composition of the segmented structure. Size (volume and maximal diameter) and shape measures (sphericity, compactness and surface-to-volume ratio) have been described in radiomics research.[37,61] Finally, features relating to tumour location (*e.g.* lung lobe, distance to the organ at risk) have been used.[72]

Image pre-processing and filtering

Most radiomics studies have used the reconstructed CT data available in the digital imaging and communications in medicine

files, without any pre-processing or normalization.[61] Possible image pre-processing includes image smoothing by average or Gaussian filters reducing image noise and image enhancement techniques such as histogram equalization, deblurring and resampling.[68,73] Image filtering has been performed, *e.g.* using a Gaussian filter followed by a Laplacian on the image, resulting in so-called Laplacian of Gaussian features.[37,74]

Computation of textural features requires discretization of the intensity values. Aerts et al[61] used equally spaced 25-HU bins for CT data. In FDG-PET data, Tixier et al[15] described the sensitivity of several textural features to varying discretization values, while Leijenaar et al[22] have shown that a fixed bin resolution enhances the interpatient and intrapatient reproducibility of feature values.

Software implementation

Several open-source software packages are available to conduct radiomics analyses. The "imaging biomarker explorer" was specifically designed to facilitate collaborations in radiomics research, offering standardized settings for image pre-processing and feature extraction.[73] This addresses the important need for benchmarking in the rapidly growing radiomics field.[75] The "Chang Gung Image Texture Analysis"[76] offers FDG-PET image analysis, while the MaZda package was originally designed for MR images.[77] The commercial packages RADIOMICS™ (OncoRadiomics, Maastricht, Netherlands) and TexRAD™ (Feedback plc, Cambridge, UK) have been developed. An overview of the main functionalities of these packages is presented in Table 1. Several in-house implementations exist as well, *e.g.* coded in MATLAB® (The Mathworks®, Natick, MA) within the Computational Environment for Radiotherapy Research environment.[37,78]

VALIDATION AND MODEL BUILDING

Feature selection methods

Typically, the number of calculated image features is much larger than the sample size of patients studied. Dimensionality reduction is therefore crucial to reduce the risk of overfitting by focusing the attention of subsequent classification efforts on a subset of relevant features. As most of the features, being based on the same matrix or quantities, will exhibit some correlation with each other, intelligent feature selection strategies are required.

Filter-based selection techniques of the univariate type (*e.g.* Relief,[79,80] Fisher score and Wilcoxon) select informative features, while multivariate filter techniques (*e.g.* minimum redundancy maximum relevance[37] and mutual information) take redundancy into account as well. In a study comparing 14 filter-based feature selection techniques in combination with 18 machine-learning (ML) classification procedures in radiomics cohorts, Wilcoxon test-based feature selection showed the highest stability against data perturbation in combination with most classifiers.[81] Wrapper selection techniques are tailored to a specific classification algorithm with the algorithm being part of the subset evaluation.[80,82]

Principal component analysis performs a transformation for dimensionality reduction and can highlight outliers.[72,83] Aerts et al[61] have used the single most informative of the four

Table 1. Overview of currently available software packages for radiomics analysis (August 2016)

Software package	Imaging modality and format	ROI definition	Features and image pre-processing	Model building	Website
IBEX (free open source)	CT, PET, MR DICOM, Pinnacle native format	DICOM-RT Editing and free drawing	109: intensity, texture, shape Smoothing, resampling, enhancement	Validation of existing models	http://bit.ly/ IBEX_MDAnderson
CGITA (free open source)	Designed for PET; CT, MR tested-DICOM	DICOM-RT, PMOD Region growing and FCM	72: intensity, texture No pre-processing	None	http://code.google. com/p/cgita
MaZda (free open source)	Designed for MR DICOM	Thresholding, deformable surface	279: intensity, texture, shape, wavelet Resampling, discretization, normalization	ANN, clustering Fisher, MI, PCA	http://www.eletel.p. lodz.pl/ programy/mazda/
RADIOMICS™ (OncoRadiomics, Maastricht, Netherlands) (commercial)	CT, PET, MR DICOM	DICOM-RT Plug-in for several TPS	543: intensity, texture, shape, wavelet Laplacian of Gaussian Resampling, discretization	None	http://www. oncoradiomics.com
TexRAD™ (Feedback plc, Cambridge, UK) (commercial)	CT, PET, MR DICOM	DICOM-RT Editing, thresholding	−30: texture and filtering (Laplacian of Gaussian)	Data-mining tool	http://www.texrad.com

ANN, Artificial Neural Network; CGITA, Chang Gung Image Texture Analysis; DICOM, digital imaging and communications in medicine; DICOM-RT, Digital Imaging and Communications in Medicine-Radiation Therapy; FCM, fuzzy c-means; IBEX, imaging biomarker explorer; MI, Mutual Information; PCA, Principal Component Analysis; PET, positron emission tomography; PMOD, PMOD Technologies LLC; ROI, region of interest; TPS, Treatment Planning System. Their characteristics and functionalities for the four main steps of the radiomics workflow are summarized.

studied feature categories to reduce redundancy. The false discovery rate[37,84] has been used to increase testing power compared with the conservative Bonferroni correction. The dynamic range of feature values can also be an important selection criterion.[72]

Model training and validation
Ultimately, the goal of most radiomics analyses is to obtain a prognostic or predictive model with a high accuracy and efficiency. Ideally, enough data are available to train and validate a classification model while holding out part of the data for the assessment of accuracy (external validation).

Unsupervised ML analyses using heat maps or clustering summarize feature data without involving an outcome variable.[85] Supervised ML classifiers such as generalized linear models, random forests,[86] support vector machines[87] and neural networks separate the data with respect to an outcome variable. Among the 18 supervised ML classification procedures studied by Parmar et al,[81] random forests had the highest performance in combination with most feature selection techniques. Moreover, the choice of classification method was found to be the dominant source of performance variation. The number of features selected as input to the classification models had a negligible impact, explaining <2% of the total variance.

Regularization methods control the complexity of a prediction model to prevent overfitting. The Lasso technique, using L1 regularization, was applied in combination with logistic regression and Cox regression (optimizing time to outcome).[88,89] It presents a weak dependence on irrelevant features, in contrast with L2 regularized logistic regression (and also with support vector machines and neural networks), which shows in worst case a linear relation between the number of irrelevant features and the required sample size.[90]

Cross-validation has frequently been used as internal validation with the AUC of the receiver-operating characteristic curve (or its generalization, concordance index) evaluating the performance accuracy.[7,37,91]

Finally, the understandability of classification models varies significantly and should be considered,[72] together with qualitative reporting.[92]

Integration of non-imaging variables
Conventional image-derived features such as tumour volume or diameter and maximum SUV have been used as reference models in radiomics studies.[61] Incorporating the most informative radiomic features into these reference models provides an indication of the gain in model performance. Patient and clinical characteristics might be added to the input variable list, as they potentially influence not only the outcome variable, but also the extracted radiomic features themselves.[72,89]

FUTURE OUTLOOK

Although the concept of using quantitative parameters describing imaging data is not new, the combined application in the radiomics workflow is currently under investigation by many research groups and has shown to be of added value. This renewed interest is mainly driven by the increased digitalization in the hospital, introduction of electronic medical records and easier access to large amounts of information through the hospital picture archiving and communication systems combined with the increased computational power. In this era of Big Data, in medical research fields also, a new field of clinical data science is emerging that allows the integration of multiple data sources to personalize treatment.[93] Radiomics is expected to be a critical component for the integration of image-derived information in this framework. Currently, imaging research is typically performed retrospectively on data stored in the picture archiving and communication systems. As described in this review, an investigation on the quality of the input imaging data is necessary to assess the potential influence on outcome prediction accuracy.

There is, however, a concern related to the use of image-derived features as such in the individualization of patient care. Treatment interventions need to be optimized and frequently (not always) this is performed on a biological or methodological basis where two treatments are compared with each other. Radiomics quantifies the phenotype of the underlying biology; however, this link is not straightforward.[94,95] More research is needed into the correlation between the underlying biological processes and the perceived image data sets. If this link is established, treatment can be tailored to the individual patient based on the imaged properties of tumours. This integration with biology might then also allow for hypothesis-driven research that is still necessary to improve the methodological understanding of the biological processes. Furthermore, a reliable predictive or prognostic power is necessary if implemented in clinical routine for individual decision-making.

FUNDING

REFERENCES

1. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009; **45**: 228–47. doi: https://doi.org/10.1016/j.ejca.2008.10.026

2. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med* 2009; **50**(Suppl 1): 122S–50S. doi: https://doi.org/10.2967/jnumed.108.057307

3. de Bruijne M. Machine learning approaches in medical image analysis: from detection to diagnosis. *Med Image Anal* 2016; **33**: 94–7. doi: https://doi.org/10.1016/j.media.2016.06.032

4. Abe Y, Hanai K, Nakano M, Ohkubo Y, Hasizume T, Kakizaki T, et al. A computer-aided diagnosis (CAD) system in lung cancer screening with computed tomography. *Anticancer Res* 2005; **25**: 483–8.

5. Li F. Potential clinical impact of advanced imaging and computer-aided diagnosis in chest radiology: importance of radiologist's role and successful observer study. *Radiol Phys Technol* 2015; **8**: 161–73. doi: https://doi.org/10.1007/s12194-015-0319-0

6. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012; **48**: 441–6. doi: https://doi.org/10.1016/j.ejca.2011.11.036

7. Fried DV, Tucker SL, Zhou S, Liao Z, Mawlawi O, Ibbott G, et al. Prognostic value and reproducibility of pretreatment CT texture features in stage III non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 2014; **90**: 834–42. doi: https://doi.org/10.1016/j.ijrobp.2014.07.020

8. Balagurunathan Y, Kumar V, Gu Y, Kim J, Wang H, Liu Y, et al. Test–retest reproducibility analysis of lung CT image features. *J Digit Imaging* 2014; **27**: 805–23. doi: https://doi.org/10.1007/s10278-014-9716-x

9. 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: https://doi.org/10.1097/rli.0000000000000180

10. 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**: 23428. doi: https://doi.org/10.1038/srep23428

11. Zhao B, James LP, Moskowitz CS, Guo P, Ginsberg MS, Lefkowitz RA, et al. Evaluating variability in tumor measurements from same-day repeat CT scans of patients with non-small cell lung cancer. *Radiology* 2009; **252**: 263–72. doi: https://doi.org/10.1148/radiol.2522081593

12. Fave X, Mackin D, Yang J, Zhang J, Fried D, Balter P, et al. Can radiomics features be reproducibly measured from CBCT images for patients with non-small cell lung cancer? *Med Phys* 2015; **42**: 6784–97. doi: https://doi.org/10.1118/1.4934826

13. Oliver JA, Budzevich M, Zhang GG, Dilling TJ, Latifi K, Moros EG. Variability of image features computed from conventional and respiratory-gated PET/CT images of lung cancer. *Transl Oncol* 2015; **8**: 524–34. doi: https://doi.org/10.1016/j.tranon.2015.11.013

14. Leijenaar RT, Carvalho S, Velazquez ER, van Elmpt WJ, Parmar C, Hoekstra OS, et al. Stability of FDG-PET radiomics features: an integrated analysis of test-retest and inter-observer variability. *Acta Oncol* 2013; **52**: 1391–7. doi: https://doi.org/10.3109/0284186X.2013.812798

15. Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. *J Nucl Med* 2012; **53**: 693–700. doi: https://doi.org/10.2967/jnumed.111.099127

16. Galavis PE, Hollensen C, Jallow N, Paliwal B, Jeraj R. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol* 2010; **49**: 1012–6. doi: https://doi.org/10.3109/0284186x.2010.498437

17. van Velden FH, Kramer GM, Frings V, Nissen IA, Mulder ER, de Langen AJ, et al. Repeatability of radiomic features in non-small-cell lung cancer [(18)F]FDG-PET/CT studies: impact of reconstruction and delineation.

Mol Imaging Biol 2016; **18**: 788–95. doi: https://doi.org/10.1007/s11307-016-0940-2

18. Yan J, Chu-Shern JL, Loi HY, Khor LK, Sinha AK, Quek ST, et al. Impact of image reconstruction settings on texture features in 18F-FDG PET. *J Nucl Med* 2015; **56**: 1667–73. doi: https://doi.org/10.2967/jnumed.115.156927

19. Yip S, McCall K, Aristophanous M, Chen AB, Aerts HJ, Berbeco R. Comparison of texture features derived from static and respiratory-gated PET images in non-small cell lung cancer. *PLoS One* 2014; **9**: e115510. doi: https://doi.org/10.1371/journal.pone.0115510

20. Grootjans W, Tixier F, van der Vos CS, Vriens D, Le Rest CC, Bussink J, et al. The impact of optimal respiratory gating and image noise on evaluation of intratumor heterogeneity on 18F-FDG PET imaging of lung cancer. *J Nucl Med* 2016; **57**: 1692–8.

21. Doumou G, Siddique M, Tsoumpas C, Goh V, Cook GJ. The precision of textural analysis in (18)F-FDG-PET scans of oesophageal cancer. *Eur Radiol* 2015; **25**: 2805–12. doi: https://doi.org/10.1007/s00330-015-3681-8

22. Leijenaar RT, Nalbantov G, Carvalho S, van Elmpt WJ, Troost EG, Boellaard R, et al. The effect of SUV discretization in quantitative FDG-PET radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep* 2015; **5**: 11075. doi: https://doi.org/10.1038/srep11075

23. Lu L, Lv W, Jiang J, Ma J, Feng Q, Rahmim A, et al. Robustness of radiomic features in [11C]choline and [18F]FDG PET/CT imaging of nasopharyngeal carcinoma: impact of segmentation and discretization. *Mol Imaging Biol* 2016; **18**: 935–45.

24. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al; European Association of Nuclear Medicine (EANM). FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging* 2015; **42**: 328–54. doi: https://doi.org/10.1007/s00259-014-2961-x

25. Wibmer A, Hricak H, Gondo T, Matsumoto K, Veeraraghavan H, Fehr D, et al. Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores. *Eur Radiol* 2015; **25**: 2840–50. doi: https://doi.org/10.1007/s00330-015-3701-8

26. Gnep K, Fargeas A, Gutiérrez-Carvajal RE, Commandeur F, Mathieu R, Ospina JD, et al. Haralick textural features on T2-weighted MRI are associated with biochemical recurrence following radiotherapy for peripheral zone prostate cancer. *J Magn Reson Imaging* 2017; **45**; 103–17. doi: https://doi.org/10.1002/jmri.25335.

27. Weygand J, Fuller CD, Ibbott GS, Mohamed ASR, Ding Y, Yang J, et al. Spatial precision in magnetic resonance imaging-guided radiation therapy: the role of geometric distortion. *Int J Radiat Oncol Biol Phys* 2016; **95**: 1304–16. doi: https://doi.org/10.1016/j.ijrobp.2016.02.059

28. Sudarshan VK, Mookiah MR, Acharya UR, Chandran V, Molinari F, Fujita H, et al. Application of wavelet techniques for cancer diagnosis using ultrasound images: a review. *Comput Biol Med* 2016; **69**: 97–111. doi: https://doi.org/10.1016/j.compbiomed.2015.12.006

29. Andrėkutė K, Linkeviciutė G, Raišutis R, Valiukevicienė S, Makštienė J. Automatic differential diagnosis of melanocytic skin tumors using ultrasound data. *Ultrasound Med Biol* 2016; **42**: 2834–43. doi: https://doi.org/10.1016/j.ultrasmedbio.2016.07.026

30. 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**: 35632. doi: https://doi.org/10.1038/srep35632

31. Song G, Xue F, Zhang C. A model using texture features to differentiate the nature of thyroid nodules on sonography. *J Ultrasound Med* 2015; **34**: 1753–60. doi: https://doi.org/10.7863/ultra.15.14.10045

32. 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. doi: https://doi.org/10.1118/1.4921123

33. Ardakani AA, Gharbali A, Mohammadi A. Classification of breast tumors using sonographic texture analysis. *J Ultrasound Med* 2015; **34**: 225–31. doi: https://doi.org/10.7863/ultra.34.2.225

34. Nadeau MJ, Desrochers A, Lamontagne M, Lariviere C, Gagnon DH. Quantitative ultrasound imaging of Achilles tendon integrity in symptomatic and asymptomatic individuals: reliability and minimal detectable change. *J Foot Ankle Res* 2016; **9**: 30. doi: https://doi.org/10.1186/s13047-016-0164-3

35. Cook GJ, Yip C, Siddique M, Goh V, Chicklore S, Roy A, et al. Are pretreatment 18F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? *J Nucl Med* 2013; **54**: 19–26. doi: https://doi.org/10.2967/jnumed.112.107375

36. Coroller TP, Agrawal V, Narayan V, Hou Y, Grossmann P, Lee SW, et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer. *Radiother Oncol* 2015; **119**: 480–6. doi: https://doi.org/10.1016/j.radonc.2016.04.004

37. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RT, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiother Oncol* 2015; **114**: 345–50. doi: https://doi.org/10.1016/j.radonc.2015.02.015

38. Desseroit MC, Visvikis D, Tixier F, Majdoub M, Perdrisot R, Guillevin R, et al. Development of a nomogram combining clinical staging with (18)F-FDG PET/CT image features in non-small-cell lung cancer stage I–III. *Eur J Nucl Med Mol Imaging* 2016; **43**: 1477–85. doi: https://doi.org/10.1007/s00259-016-3325-5

39. Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, et al. Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. *Radiology* 2016; **281**: 947–57.

40. Liang C, Huang Y, He L, Chen X, Ma Z, Dong D, et al. The development and validation of a CT-based radiomics signature for the preoperative discrimination of stage I–II and stage III–IV colorectal cancer. *Oncotarget* 2016; **7**: 31401–12. doi: https://doi.org/10.18632/oncotarget.8919

41. Sullivan DC, Obuchowski NA, Kessler LG, Raunig DL, Gatsonis C, Huang EP, et al. Metrology standards for quantitative imaging biomarkers. *Radiology* 2015; **277**: 813–25. doi: https://doi.org/10.1148/radiol.2015142202

42. Buckler AJ, Bresolin L, Dunnick NR, Sullivan DC; Group. A collaborative enterprise for multi-stakeholder participation in the advancement of quantitative imaging. *Radiology* 2011; **258**: 906–14. doi: https://doi.org/10.1148/radiol.10100799

43. Eminowicz G, McCormack M. Variability of clinical target volume delineation for definitive radiotherapy in cervix cancer. *Radiother Oncol* 2015; **117**: 542–7. doi: https://doi.org/10.1016/j.radonc.2015.10.007

44. Echegaray S, Gevaert O, Shah R, Kamaya A, Louie J, Kothary N, et al. Core samples for radiomics features that are insensitive to tumor segmentation: method and pilot study using CT images of hepatocellular carcinoma. *J Med Imaging (Bellingham)* 2015; **2**: 041011. doi: https://doi.org/10.1117/1.JMI.2.4.041011

45. 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: https://doi.org/10.1593/tlo.13844

46. Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al.

Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS One* 2014; **9**: e102107. doi: https://doi.org/10.1371/journal.pone.0102107

47. Velazquez ER, Parmar C, Jermoumi M, Mak RH, van Baardwijk A, Fennessy FM, et al. Volumetric CT-based segmentation of NSCLC using 3D-slicer. *Sci Rep* 2013; **3**: 3529. doi: https://doi.org/10.1038/srep03529

48. van Baardwijk A, Bosmans G, Boersma L, Buijsen J, Wanders S, Hochstenbag M, et al. PET-CT-based auto-contouring in non-small-cell lung cancer correlates with pathology and reduces interobserver variability in the delineation of the primary tumor and involved nodal volumes. *Int J Radiat Oncol Biol Phys* 2007; **68**: 771–78. doi: https://doi.org/10.1016/j.ijrobp.2006.12.067

49. Bezdek JC, Ehrlich R, Full W. FCM: The fuzzy c-means clustering algorithm. *Comput Geosci* 1984; **10**: 191–203.

50. Hatt M, Lamare F, Boussion N, Turzo A, Collet C, Salzenstein F, et al. Fuzzy hidden Markov chains segmentation for volume determination and quantitation in PET. *Phys Med Biol* 2007; **52**: 3467–91. doi: https://doi.org/10.1088/0031-9155/52/12/010

51. Hatt M, Cheze le Rest C, Turzo A, Roux C, Visvikis D. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. *IEEE Trans Med Imaging* 2009; **28**: 881–93. doi: https://doi.org/10.1109/TMI.2008.2012036

52. Hatt M, Cheze le Rest C, Descourt P, Dekker A, De Ruysscher D, Oellers M, et al. Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications. *Int J Radiat Oncol Biol Phys* 2010; **77**: 301–8. doi: https://doi.org/10.1016/j.ijrobp.2009.08.018

53. Hatt M, Cheze-le Rest C, van Baardwijk A, Lambin P, Pradier O, Visvikis D. Impact of tumor size and tracer uptake heterogeneity in (18)F-FDG PET and CT non-small cell lung cancer tumor delineation. *J Nucl Med* 2011; **52**: 1690–7. doi: https://doi.org/10.2967/jnumed.111.092767

54. Hatt M, Tixier F, Cheze Le Rest C, Pradier O, Visvikis D. Robustness of intratumour ^{18}F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *Eur J Nucl Med Mol Imaging* 2013; **40**: 1662–71. doi: https://doi.org/10.1007/s00259-013-2486-8

55. Sorensen L, Nielsen M, Lo P, Ashraf H, Pedersen JH, de Bruijne M. Texture-based analysis of COPD: a data-driven approach. *IEEE Trans Med Imaging* 2012; **31**: 70–8. doi: https://doi.org/10.1109/TMI.2011.2164931

56. National Lung Screening Trial Research Team; Aberle DR, Berg CD, Black WC, Church TR, Fagerstrom RM, Galen B, et al. The National Lung Screening Trial: overview and study design. *Radiology* 2011; **258**: 243–53. doi: https://doi.org/10.1148/radiol.10091808

57. Hawkins S, Wang H, Liu Y, Garcia A, Stringfield O, Krewer H, et al. Predicting malignant nodules from screening CT scans. *J Thorac Oncol* 2016; **11**: 2120–8. doi: https://doi.org/10.1016/j.jtho.2016.07.002

58. Huang K, Senthi S, Palma DA, Spoelstra FO, Warner A, Slotman BJ, et al. High-risk CT features for detection of local recurrence after stereotactic ablative radiotherapy for lung cancer. *Radiother Oncol* 2013; **109**: 51–7. doi: https://doi.org/10.1016/j.radonc.2013.06.047

59. Peulen H, Mantel F, Guckenberger M, Belderbos J, Werner-Wasik M, Hope A, et al. Validation of high-risk computed tomography features for detection of local recurrence after stereotactic body radiation therapy for early-stage non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 2016; **96**: 134–41. doi: https://doi.org/10.1016/j.ijrobp.2016.04.003

60. Cunliffe A, Armato SG 3rd, 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: https://doi.org/10.1016/j.ijrobp.2014.11.030

61. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014; **5**: 4006. doi: https://doi.org/10.1038/ncomms5006

62. van Velden FH, Cheebsumon P, Yaqub M, Smit EF, Hoekstra OS, Lammertsma AA, et al. Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous intratumoural FDG uptake in non-small cell lung cancer PET studies. *Eur J Nucl Med Mol Imaging* 2011; **38**: 1636–47. doi: https://doi.org/10.1007/s00259-011-1845-6

63. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cyb* 1973; **3**: 610–21. doi: https://doi.org/10.1109/tsmc.1973.4309314

64. Amadasun M, King R. Textural features corresponding to textural properties. *IEEE Trans Syst Man Cyb* 1989; **19**: 1264–74. doi: https://doi.org/10.1109/21.44046

65. Sun C, Wee WG. Neighboring gray level dependence matrix for texture classification. *Comput Vision Graph* 1982; **23**: 341–52. doi: https://doi.org/10.1016/0734-189x(83)90032-4

66. Galloway MM. Texture analysis using gray level run lengths. *Comput Vision Graph* 1975; **4**: 172–9. doi: https://doi.org/10.1016/s0146-664x(75)80008-6

67. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges JP, 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. doi: https://doi.org/10.2967/jnumed.110.082404

68. Vallieres M, Freeman CR, Skamene SR, El Naqa I. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys Med Biol* 2015; **60**: 5471–96. doi: https://doi.org/10.1088/0031-9155/60/14/5471

69. Ganeshan B, Goh V, Mandeville HC, Ng QS, Hoskin PJ, Miles KA. Non-small cell lung cancer: histopathologic correlates for texture parameters at CT. *Radiology* 2013; **266**: 326–36. doi: https://doi.org/10.1148/radiol.12112428

70. Ng F, Kozarski R, Ganeshan B, Goh V. Assessment of tumor heterogeneity by CT texture analysis: can the largest cross-sectional area be used as an alternative to whole tumor analysis? *Eur J Radiol* 2013; **82**: 342–8. doi: https://doi.org/10.1016/j.ejrad.2012.10.023

71. Fave X, Cook M, Frederick A, Zhang L, Yang J, Fried D, et al. Preliminary investigation into sources of uncertainty in quantitative imaging features. *Comput Med Imaging Graph* 2015; **44**: 54–61. doi: https://doi.org/10.1016/j.compmedimag.2015.04.006

72. 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: https://doi.org/10.1016/j.mri.2012.06.010

73. 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: https://doi.org/10.1118/1.4908210

74. Ganeshan B, Miles KA. Quantifying tumour heterogeneity with CT. *Cancer Imaging* 2013; **13**: 140–9. doi: https://doi.org/10.1102/1470-7330.2013.0015

75. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016; **278**: 563–77. doi: https://doi.org/10.1148/radiol.2015151169

76. Fang YH, Lin CY, Shih MJ, Wang HM, Ho TY, Liao CT, et al. Development and evaluation of an open-source software package "CGITA" for quantifying tumor

heterogeneity with molecular images. *Biomed Res Int* 2014; **2014**: 248505. doi: https://doi.org/10.1155/2014/248505

77. Szczypinski PM, Strzelecki M, Materka A, Klepaczko A. MaZda—a software package for image texture analysis. *Comput Methods Programs Biomed* 2009; **94**: 66–76. doi: https://doi.org/10.1016/j.cmpb.2008.08.005

78. Deasy JO, Blanco AI, Clark VH. CERR: a computational environment for radiotherapy research. *Med Phys* 2003; **30**: 979–85. doi: https://doi.org/10.1118/1.1568978

79. Kira K, Rendell LA. The feature selection problem: traditional methods and a new algorithm. AAAI-92 proceedings; 1992. pp. 129–34.

80. Basu S, Hall LO, Goldgof DB, Gu Y, Kumar V, Choi J, et al. Developing a classifier model for lung tumours in CT-scan images. *IEEE Trans Syst Man Cyb* 2011: 1306–12.

81. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJ. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep* 2015; **5**: 13087. doi: https://doi.org/10.1038/srep13087

82. Kohavi R, John GH. Wrappers for feature subset selection. *Artif Intell* 1997; **97**: 273–324. doi: https://doi.org/10.1016/s0004-3702(97)00043-x

83. Song F, Guo Z, Mei D. Feature selection using principal component analysis. *International Conference System Science, Engineering Design Manufacturing Informatization*; 2010.

84. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 1995; **57**: 289–300.

85. Parmar C, Leijenaar RT, Grossmann P, Rios Velazquez E, Bussink J, Rietveld D, et al. Radiomic feature clusters and prognostic signatures specific for lung and head & neck cancer. *Sci Rep* 2015; **5**: 11044. doi: https://doi.org/10.1038/srep11044

86. Hoeben BA, Starmans MH, Leijenaar RT, Dubois LJ, van der Kogel AJ, Kaanders JH, et al. Systematic analysis of 18F-FDG PET and metabolism, proliferation and hypoxia markers for classification of head and neck tumors. *BMC Cancer* 2014; **14**: 130. doi: https://doi.org/10.1186/1471-2407-14-130

87. Gao X, Chu C, Li Y, Lu P, Wang W, Liu W, et al. The method and efficacy of support vector machine classifiers based on texture features and multi-resolution histogram from (18)F-FDG PET-CT images for the evaluation of mediastinal lymph nodes in patients with lung cancer. *Eur J Radiol* 2015; **84**: 312–7. doi: https://doi.org/10.1016/j.ejrad.2014.11.006

88. Tibshirani R. Regression shrinkage and selection *via* the Lasso. *J R Stat Soc Series B Stat Methodol* 1996; **58**: 267–88.

89. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol* 2016; **34**: 2157–64. doi: https://doi.org/10.1200/JCO.2015.65.9128

90. Ng AY. Feature selection, L1 *vs* L2 regularization, and rotational invariance. *Proceedings of the 21st International Conference on Machine Learning*; 2004.

91. Parmar C, Grossmann P, Rietveld D, Rietbergen MM, Lambin P, Aerts HJ. Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer. *Front Oncol* 2015; **5**: 272. doi: https://doi.org/10.3389/fonc.2015.00272

92. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC Med* 2015; **13**: 1. doi: https://doi.org/10.1186/s12916-014-0241-z

93. Benedict SH, El Naqa I, Klein EE. Introduction to big data in radiation oncology: exploring opportunities for research, quality assessment, and clinical care. *Int J Radiat Oncol Biol Phys* 2016; **95**: 871–2. doi: https://doi.org/10.1016/j.ijrobp.2015.12.358

94. Panth KM, Leijenaar RT, Carvalho S, Lieuwes NG, Yaromina A, Dubois L, et al. Is there a causal relationship between genetic changes and radiomics-based image features? An *in vivo* preclinical experiment with doxycycline inducible GADD34 tumor cells. *Radiother Oncol* 2015; **116**: 462–6. doi: https://doi.org/10.1016/j.radonc.2015.06.013

95. Segal E, Sirlin CB, Ooi C, Adler AS, Gollub J, Chen X, et al. Decoding global gene expression programs in liver cancer by noninvasive imaging. *Nat Biotechnol* 2007; **25**: 675–80. doi: https://doi.org/10.1038/nbt1306