BJR

Received: 25 May 2014 | Revised: 22 July 2014 | Accepted: 22 July 2014

doi: 10.1259/bjr.20140369

Cite this article as:
Alobaidli S, McQuaid S, South C, Prakash V, Evans P, Nisbet A. The role of texture analysis in imaging as an outcome predictor and potential tool in radiotherapy treatment planning. Br J Radiol 2014;87:20140369.

REVIEW ARTICLE

The role of texture analysis in imaging as an outcome predictor and potential tool in radiotherapy treatment planning

[1]S ALOBAIDLI, MSc, [2]S MCQUAID, PhD, [2]C SOUTH, PhD, [3]V PRAKASH, MB ChB, FRCP,' [1]P EVANS, DPhil, and [2,4]A NISBET, PhD

[1]Centre for Vision, Speech and Signal Processing, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, UK
[2]Department of Medical Physics, Royal Surrey County Hospital NHS Foundation Trust, Guildford, UK
[3]Department of Radiology, Ashford and St Peter's Hospital NHS Foundation Trust, Ashford, UK
[4]Centre for Nuclear and Radiation Physics, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, UK

Address correspondence to: Professor Andrew Nisbet
E-mail: *andrew.nisbet@nhs.net*

ABSTRACT

Predicting a tumour's response to radiotherapy prior to the start of treatment could enhance clinical care management by enabling the personalization of treatment plans based on predicted outcome. In recent years, there has been accumulating evidence relating tumour texture to patient survival and response to treatment. Tumour texture could be measured from medical images that provide a non-invasive method of capturing intratumoural heterogeneity and hence could potentially enable a prior assessment of a patient's predicted response to treatment. In this article, work presented in the literature regarding texture analysis in radiotherapy in relation to survival and outcome is discussed. Challenges facing integrating texture analysis in radiotherapy planning are highlighted and recommendations for future directions in research are suggested.

Radiotherapy has seen rapid development in the past few decades with delineation of target volumes becoming more precise, owing to improvements in medical imaging modalities.[1] In the field of radiation oncology, medical images are used to diagnose, stage, plan and assess the response of tumours to treatment.[1] Today, using CT to guide radiotherapy planning is considered to be the standard of care. In addition to its high image resolution (approximately 1 mm), radiotherapy treatment planning systems rely on Hounsfield units obtained from CT for dose calculation. Yet, CT-based radiotherapy planning relies on qualitative information from CT images, where quantitative measurements are usually confined to determining tumour size in three dimensions. Even though tumours have been shown to be biologically heterogeneous,[2] radiotherapy planning is still often based on TNM staging. The TNM system is used to stage cancer: a letter and a number are assigned to each letter in the TNM system, where T stands for the tumour, referring to its size; N stands for node and represents the status of lymph node involvement; and M is for metastasis, providing information on whether the cancer has spread to other parts of the body.[3] Given the increased evidence of intratumoural heterogeneity and the observed diversity of patients' response to treatment, the field of oncology has been striving for personalized treatment based on prognostic information of a tumour's response to treatment.

In recent years, nuclear medicine imaging, especially positron emission tomography (PET) has been increasingly utilized in the field of oncology as a source of quantitative measurements regarding tumour biology. Fluorine-18 fludeoxyglucose (^{18}F-FDG) PET imaging has probably made a greater contribution to radiotherapy planning with regard to assessing response to treatment than have other functional imaging modalities such as functional MRI (fMRI). The wide availability, high sensitivity and the relatively well-established protocols and quantification methods of ^{18}F-FDG PET have led to its subsequent popularity in the field of oncology.[4] Even though several studies have reported ^{18}F-FDG PET to be a predictor for a tumour's response to treatment,[5–14] the standardized uptake value (SUV) remains semi-quantitative, challenged by changes in tumour volume and elapsed time, which affects the kinetics of the uptake and subsequently the distribution of the tracer.[15,16]

Limitations in existing imaging modalities and the concept that radiological images hold more information than is being utilized has led to increased interest in the field of radiomics. Radiomics refers to quantitative feature extraction from radiological images and their use to generate meaningful data.[17] In cancer therapy, assessing tumour heterogeneity in relation to survival by extracting textural features has emerged in the past few years as a potentially useful tool. Algorithms have been proposed in the literature to increase the predictive power of existing imaging modalities by extracting tumour textural features based on intensity values. Even though the proposed methodologies are still in their early stages of development, suffering from expected shortcomings and faced with various challenges, the presented results are promising. The benefits of employing a post-processing algorithm on routinely acquired patient images to predict response to treatment, if established to be valid, would be numerous. Predicting tumours' response to therapy could lead to less aggressive treatments for patients expected to respond well, thus reducing side effects, and adapting a more rigorous treatment course for patients expected to respond poorly. The proposed texture analysis approaches in radiotherapy presented in the literature are explored in this article highlighting areas of significance, challenges and potential applications.

TEXTURE ANALYSIS

Texture of an image can be generally defined as the spatial variation in pixel intensity levels. Texture can be assessed using statistical methods (histogram and grey-level dependence matrices), model-based methods, such as fractal models, or transform-based methods, such as the Fourier and the wavelet transform.[18] In medical imaging, statistical-based texture analysis has made the most significant contribution in predicting response for patients receiving radiotherapy. Fractal analysis has been employed to assess response to chemotherapy, but its significance to response to radiotherapy has not been reported in the literature.[19,20]

Intensity histograms are a representation of the distribution of pixel intensities in an image, and they are considered first-order statistical methods where global features such as mean and standard deviation (SD) can be extracted. Second-order statistics textural features are extracted by the analysis of grey-level co-occurrence matrices (GLCM).[18,21] GLCM represents the joint probability density function $P(i, j; \theta, d)$ of the number of times an intensity level i and an intensity level j occur in a certain direction with $\theta = 0°$, 45°, 90° or 135° at specified distance d.[22] Where p_x is the probability matrix obtained from summing the rows in $p(i, j)$; p_y is the probability matrix from summing the columns in $p(i, j)$; and μ_x, μ_y, σ_x and σ_y are means and SD of p_x and p_y.

Textural parameters extracted from neighbourhood grey-tone (intensity) difference matrices (NGTDM) are considered higher order parameters, which are based on the relation between a pixel and the neighbouring pixels. Run length matrices (RLNM) are considered higher order statistical methods, where RLNM consist the number of consecutive pixels that have the same intensity level and which occur in a specified direction.[23] GLCM, NGTDM and RLNM are grey-level dependence matrices that could be generated in four directions of θ as seen in Figure 1.

In texture analysis, the order of the extracted feature refers to the relationship between the pixels of the extracted features. First-order statistics are extracted from the average of pixel intensities, and second-order statistical features are measured based on the relationship between two pixels, where the higher order statistical features are based on the relationship between more than two pixels as illustrated in Figure 2. Therefore, first-order statistical methods do not convey spatial information, whereas the second and higher order statistical methods maintain spatial information. Table 1 summarizes statistical textural features that can be extracted.[18,22–24]

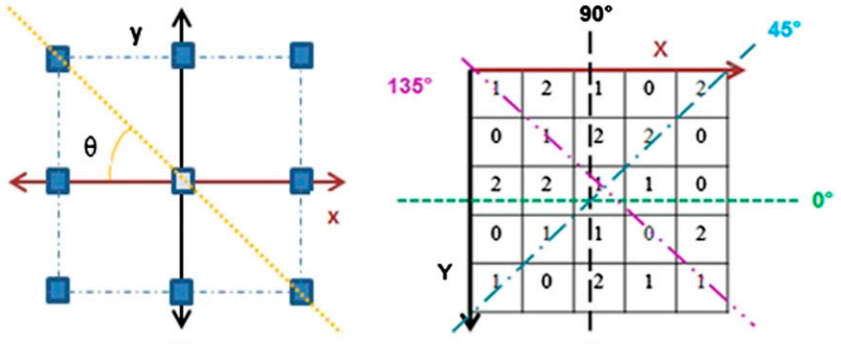
Textural features such as mean, variance, entropy, energy, contrast and coarseness are well understood with regard to the information that they provide of an image's textural properties and have been heavily used to quantify image texture in the literature. Yet, the information extracted about image texture properties from other features, such as sum entropy, sum variance and information measures of correlation are not clear with regard to the clinical aspect. However, these features are presented in Table 1, as they could prove to be clinically important in the future and to provide the complete set of statistical texture measures.

Table 1 presents a full set of the textural features that could be measured from first-order, second-order and third-order statistics. N is the number of grey levels in an image, and i is the intensity level of a pixel. For the first-order statistics, $p(i)$ is the probability that an intensity level i occurs in an image. Entropy extracted from first-order statistics is denoted as H, where entropy measured from second-order statistics based on the joint probability $P(i, j; \theta, d)$ and is denoted as HXY. Entropy may also be measured from the probability matrix obtained from summing the rows in $p(i, j)$ (denoted by p_x), and it is denoted as HX in Equation (18). Entropy measured from the probability matrix from summing the columns in $p(i,j)$ is denoted HY as seen in Equation (18). In Equation (10), μ_x, μ_y, σ_x and σ_y are means and SD of p_x and p_y, where $p_y(i) = \sum_{i=1}^{N} P(i,j)$ and $p_x(i) = \sum_{j=1}^{N} P(i,j)$ and $p_{x+y}(k) = \sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j)$, where $k = i + j$.

For higher order statistics, $s(i)$ is the NGTDM; G_h is the highest intensity value in the region of interest (ROI); and n is the number of pixels in ROI.

To date, textural analysis in the medical field has been used for tissue classification,[25] such as lung pathologies,[26,27] breast lesions,[28] head and neck tumours[29] and glioneural tumours.[30] Furthermore, NGTDM-based textural analysis was suggested by Yu et al[29] for automatic delineation of target volumes in head and neck tumours for radiotherapy planning. In recent years, statistical-based texture analysis has been suggested to have prognostic value in radiotherapy. Several textural features could be extracted from grey-level dependence matrices and NGTDM, yet based on reviewed literature, contrast, energy, entropy and homogeneity have shown the most significant results in relation to patients' survival.

Figure 1. The two-dimensional displacement in generating grey-level dependence matrices is shown in (a). (b) Represents a sample image showing all the four different directions from which the grey-level dependence matrices could be generated. The general form for constructing a grey dependence matrix is illustrated in (c).



j

i	0	1	2
0	#(0,0)	#(0,1)	#(0,2)
1	#(1,0)	#(1,1)	#(1,2)
2	#(2,0)	#(2,1)	#(2,2)

(c)

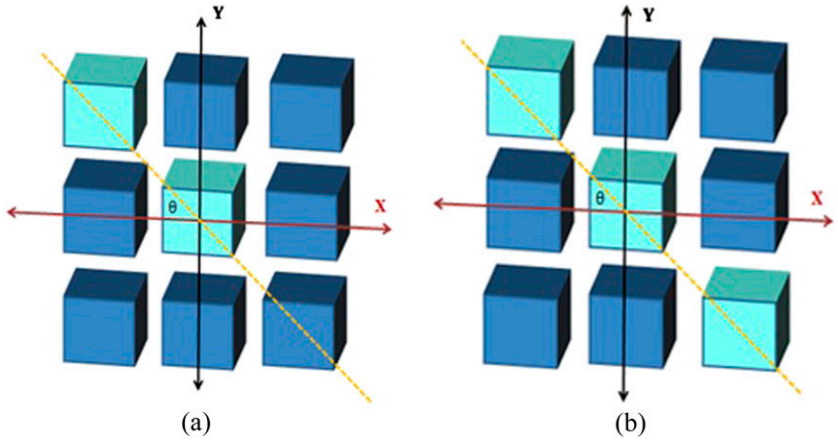
TEXTURE ANALYSIS IN CT

The initial implementation of texture analysis to predict patient survival was suggested by Ganeshan et al[31] in 2007. The authors suggested band-pass filtering of CT images using the Laplacian of gaussian (LoG) to detect features at different anatomical scales. Extracting texture parameters from CT images using a filtration–histogram method was recommended by the authors, given their uncomplex and rotationally invariant nature despite their spatial insensitivity. The authors performed texture analysis on CT images using in-house software, which has now became commercially available by the name of TexRAD (TexRAD Ltd, Cambridge, UK), and a number of studies using this software have been published.[32–39] The group suggested the use of five LoG filters ranging from fine to coarse textures (filter widths: fine = 4 image pixels and coarse = 12 image pixels) and quantifying tumour heterogeneity by measuring the mean, SD, skewness, mean value of positive pixels (MPP), uniformity and entropy from the CT slice with the largest cross section of the tumour. The group reported uniformity, which is inversely proportional to entropy, to show the most significant results in predicting patient survival.

A recent publication using TexRAD investigated whether first-order textural features extracted from staging PET/CT images, for 56 patients with non-small-cell lung carcinoma (NSCLC), were related to patient survival.[40] By quantifying the entropy of an image, which represents the pixel distribution in the image, tumour heterogeneity was assessed. To correlate the heterogeneity of the tumours to patient survival, Kaplan–Meier (KM) curves were generated. Patients were assigned to groups of low and high entropy based on the optimum threshold value obtained from a feasibility study including the same number of patients. The entropy value that showed most optimum in predicting survival was a normalized value taken from medium filter/coarse filter entropy values. The authors reported tumour normalized entropy to be the only parameter significantly related ($p = 0.027$) to survival for patients in the curative-intent group, while tumour entropy, stage and permeability showed significant association with survival for the palliative patients group with $p = 0.042$, $p = 0.020$ and $p = 0.003$, respectively. Although the published articles have shown some correlation between survival and global textural parameters, they have not explained the significance of choosing the optimum filter width and how that relates to anatomical feature extraction. Moreover, TexRad-based studies have not validated the uniformity threshold values used for generating KM curves. In addition, the proposed methodology by the TexRAD group based on using the tumour's image slice with the largest cross section is not a volumetric analysis, which leads to loss of spatial information and consequently affects textural information.

Correlation of patients' survival to second-order textural parameters extracted from CT data was investigated by Vaidya et al[41] in 2012. The study included 27 patients diagnosed with NSCLC confined to the thorax, where baseline ^{18}FDG PET and the corresponding CT were analysed. The end points were set to be local recurrence with at least 6 months follow-up. To account for patient breathing artefacts, an algorithm based on the inverse filtering process was applied to acquired images. Textural features extracted from PET/CT data included energy, contrast, entropy and local homogeneity from co-occurrence matrices calculated for the three-dimensional (3D) volume of the tumour. Textural features extracted from CT images showed some correlation with locoregional failure, yet none of the features

Figure 2. (a) An example of pixel pair relationship in generating grey-level co-occurrence matrices where (b) represents the pixel relationship in neighbourhood grey-tone (intensity) difference matrices and run length matrices.



reached significance. The deblurring method was applied to PET images not CT images, so whether motion artefacts affected the power of extracted features was not determined. The authors reported extracted textural features from PET images to have the lowest correlation with locoregional failure when compared with intensity–volume histogram (IVH) and gross tumour volume (GTV) and SUV measurements. The deblurring of PET images showed little improvement on the correlation coefficient r (from 0.114 to 0.180 and p-value from 0.28 to 0.17) for local homogeneity. Textural features showed a stronger correlation with local control rather than with locoregional failure, whereas IVH showed the opposite trend as reported by the authors.[41]

Textural features were reported by Mattonen et al[42,43] to predict recurrence for patients with early stage NSCLC who underwent stereotactic ablative radiotherapy. The authors reported that SD [reported as the variation of Hounsfield units within the ground-glass opacity (GGO) regions, which are regions where the normal lung parenchyma density is increased with visible vessels] is able to significantly discriminate ($p = 0.0078$) between patients with radiation-induced injury and patients with recurrence in a follow-up CT scan acquired at 9 months with an error of 26%.[42] The authors explored the ability of second-order textural features to predict recurrence by generating GLCM from GGO for follow-up CT scans taken at 6 months for a group of 22 patients. The texture features calculated from the two-dimensional (2D) averaged GLCM were energy, entropy, correlation, inverse difference moment, inertia (contrast), cluster shade and cluster prominence. Cluster prominence and cluster shade are measures of pixel pair grouping quantifying the symmetry (skewness in GLCM) of grey levels in an image. The authors reported energy, entropy and inertia to be significantly different between groups with radiation-induced injury and recurrence, with $p = 0.036$, $p = 0.034$ and $p = 0.036$, respectively. Respiration, comorbidities and any scanning factors that would affect CT density were not considered.[43] Given that injury to lung tissue, such as radiation fibrosis and pneumonitis, post-radiotherapy resembles tumour recurrence on CT images, a more analytical quantitative approach to predict recurrence is needed.

The recent publications on relating textural features extracted from CT images to patient survival are promising, yet need extensive validation. Factors affecting extracted textural features, such as reconstruction algorithms, scanning parameters, organ motion, number of pixels/voxels in a tumour needed to generate statistically significant textural information and whether volumetric analysis is superior to 2D analysis, are yet to be investigated. Moreover, published studies are retrospective where prospective studies are needed.

TEXTURE ANALYSIS IN POSITRON EMISSION TOMOGRAPHY

In 2009, El Naqa et al[44] investigated building a prognostic multimetric model of relevant features of pre-chemoradiotherapy ^{18}F-FDG PET accumulation in tumours. The study included 14 patients diagnosed with cervical cancer and 9 patients diagnosed with head and neck cancer. The median follow-up was 30 months. The end point for cervical cancer was disease persistence, and the end point was overall survival for head and neck cancer. Extraction of textural features followed two approaches. The first was based on the intensity histogram method termed IVH, where a single curve was generated that summarized the intensity information from the 3D functional volume for the ROI. In this study, the authors measured textural features and generated IVH for both clinical tumour volume (CTV) and GTV, where CTV was delineated by an expert oncologist including visible tumour and margins for subclinical invasion. GTV was delineated as the region within the CTV having 40% of maximum SUV for cervical cancer, and for head and neck, GTV was delineated by registering the simulation CT scan with the PET scan and then was manually delineated by the physician, given that the 40% maximum SUV is considered unreliable for head and neck tumours. The second approach was based on extracting shape features (eccentricity, Euler number, solidity and extent) and GLCM textural features (contrast, energy, homogeneity and entropy) from the GTV. The authors reported that textural features have higher predictive power than does SUV measurement, with energy showing the greatest significance with correlation coefficient $r = -0.42$ and area under the curve (AUC) = 0.72 for cervical cancer. For patients with head

Table 1. Summary of statistical texture features

Order of extracted feature	Definition	Texture feature	Equation
First-order global	Based on average pixel value. Intensity histogram analysis	Mean (μ): average intensity values in an image	$\sum_i^N ip(i)$ (1)
		Variance (σ^2): the spread or variation around the mean	$\sum_i^N (i-\mu)^2 p(i)$ (2)
		Skewness: symmetry of intensity values in an image. Skewness is zero if the histogram is symmetrical	$\sigma^{-3}\sum_i^N (i-\mu)^3 p(i)$ (3)
		Kurtosis: indication of histogram flatness	$\sigma^{-4}\sum_i^N (i-\mu)^4 p(i) - 3$ (4)
		Energy: measures uniformity of intensity values	$\sum_i^N [p(i)]^2$ (5)
		Entropy (HXY): represents irregularity of intensity value distribution	$-\sum_i^N p(i)\log[p(i)]$ (6)
Second-order local	Based on grey-level dependence matrices	Angular second moment (energy or uniformity): measures homogeneity of intensity value distribution in an image	$\sum_{i,j}^N [p(i,j)]^2$ (7)
		Contrast: measures amount of local variation in intensity values	$\sum_{i,j}^N \lvert i-j\rvert^2 p(i,j)$ (8)
		Homogeneity (inverse difference moment): measures the homogeneity of the intensity values of the pixel pair	$\sum_{i,j}^N p(i,j)/1-(i-j)^2$ (9)
		Correlation: measures the linear dependencies of intensity values in an image	$\sum_{i,j}^N (i,j)p(i,j) - \mu_x\mu_y/\sigma_x\sigma_y$ (10)
		Entropy: measure of randomness of intensity values in an image	$-\sum_{i,j}^N p(i,j)\log[p(i,j)]$ (11)
		Sum of squares (variance)	$\sum_{i,j}^N (i-\mu)^2 p(i,j)$ (12)
		Sum average	$\sum_{i=2}^{2N} ip_{x+y}(i)$ (13)
		Sum entropy	$-\sum_{i=2}^{2N} p_{x+y}(i)\log[p_{x+y}(i)]$ (14)
		Sum variance	$\sum_{i=2}^{2N} (i-\text{sum entropy})^2 p_{x+y}(i)$ (15)
		Difference variance	Variance of p_{x+y} (16)
		Difference entropy	$-\sum_{i=0}^{N-1} p_{x-y}(i)\log[p_{x-y}(i)]$ (17)
		Information measures of correlation	$\frac{\text{HXY}-\text{HXY1}}{\max(\text{HX},\text{HY})}$ (18)
			$\{1-\exp[-2.0(\text{HXY2}-\text{HXY})]\}^2$ (19) where $\text{HXY}=-\sum_{i=1}^N p(i,j)\log[p(i,j)]$ and HX and HY are entropies of p_x and p_y $\text{HXY1}=-\sum_{i=1}^N p(i,j)\log[p_x(i)p_y(j)]$ (20) $\text{HXY2}=-\sum_{i=1}^N p_x(i)p_y(j)\log[p_x(i)p_y(j)]$ (21)
		Maximal correlation coefficient	(second largest eigenvalue of $Q)^{1/2}$ (22) where $Q(i,j)=\sum_k p(i,k)p(j,k)/p_x(i)p_y(k)$ (23)

(Continued)

Table 1. (Continued)

Order of extracted feature	Definition	Texture feature	Equation
Higher order	Based on neighbourhood grey-tone difference matrices	Local	
		Complexity: measures amount of information (primitives) in texture	$\sum_{i,j}^{G_h}((\lvert i-j\rvert)/\{n^2[p(i)+p(j)]\}) \times [p(i)s(i)+p(j)s(j)]$ (24) where $s(i)$ is the neighbourhood grey-tone (intensity) difference matrices; G_h is the highest intensity value in the ROI; and n is the number of pixels in ROI
		Busyness: measures the rate of change in intensity values	$\sum_{i}^{G_h} p(i)s(i)/\sum_{i,j}^{G_h}\lvert ip(i)-jp(j)\rvert$ (25)
		Contrast: measures the variation of intensity values in an image	$[1/N(N-1)\sum_{i,j}^{G_h} p(i)p(j)(i-j)^2] \times \left[1/n\sum_{i}^{G_h} s(i)\right]$ (26)
		Coarseness: measures the density of edges in an image	$\left[\varepsilon+\sum_{i}^{G_h}\frac{s(i)}{n}\right]^{-1}$ (27) ε is a small number, so the dominator value will not be zero
		Texture strength: measures how definable (distinguishable) primitive texture is	$\{\sum_{i,j}^{G_h}[p(i)+p(j)][i-j]^2\}/[\varepsilon+\sum_{i}^{G_h} s(i)]$ (28)
		Regional	
		Short-run emphasis: measures the run length distribution emphasizing short runs by dividing by the square of run length value	$\sum_{j=1}^{N}\sum_{i=1}^{N} p(i,j)/j^2/\sum_{j=1}^{N}\sum_{i=1}^{N} p(i,j)$ (29)
		Long-run emphasis: measures the run length distribution emphasizing long runs by multiplying by the square of run length value	$\sum_{j=1}^{N}\sum_{i=1}^{N} j^2 p(i,j)/\sum_{j=1}^{N}\sum_{i=1}^{N} p(i,j)$ (30)
		Grey-level non-uniformity: represents the similarity of intensity values in an image	$\sum_{i=1}^{N}[\sum_{j=1}^{N} p(i,j)]^2/\sum_{j=1}^{N}\sum_{i=1}^{N} p(i,j)$ (31)
		Run length non-uniformity: measure the run length similarity	$\sum_{j=1}^{N}[\sum_{i=1}^{N} p(i,j)]^2/\sum_{j=1}^{N}\sum_{i=1}^{N} p(i,j)$ (32)
		Run percentage: ratio of total number of runs to the total number of possible runs measuring the homogeneity of runs. For images with most linear structure, the value of run percentage is lowest	$\sum_{i,j}^{N} p(i,j)/P$ (33)

ROI, region of interest.

and neck cancer, shape parameters showed the highest predictive power with tumour volume having the highest significance ($r = -0.6$ and AUC = 0.85). The SUV measurement was reported to have the lowest predictive value and the texture features of contrast and local homogeneity showed relatively good predicting power ($r = -0.519$; AUC = 0.8 and $r = 0.519$; AUC = 0.825, respectively). The IVH for the tumour volume having at least 90% of maximum intensity value was reported to provide the best univariate prediction for survival overall, where $r = -0.78$ and AUC = 0.95.[44]

Furthermore, GLCM-based texture analysis of ^{18}F-FDG PET data has been reported to provide significant and superior results to SUV measurements in predicting complete response

(CR), partial response (PR) and non-response (NR) to therapy, for patients diagnosed with oesophageal carcinoma as reported by Tixier et al[45] in 2011. The study included 41 patients, where SUV measurements were compared with global, regional and local PET textural features. The authors reported both global features and SUV measurements to be sufficient predictors of CR but could not distinguish NR from PR. Regional entropy was reported to be the most significant predictor in identifying and distinguishing NR, CR and PR. The images of the tumour were quantized by resampling the intensity values to the 16-, 32-, 64- or 128-set of discrete values to reduce noise and to normalize the intensities across the image. The authors assessed the variation in quantizing GLCM to discrete number of values and found no significant difference ($p < 0.05$).[45]

The reproducibility of PET textural analysis was evaluated by Tixier et al[45,46] comparing two baseline ^{18}F-FDG PET scans taken 2–7 days apart for 16 patients diagnosed with oesophageal cancer. Results showed the mean percentage difference (%DIFF) between the two studies to be 4.7% ± 19.5% and 5.5% ± 21.2% for SUV mean and SUV maximum, respectively, confirming previous studies. Local heterogeneity parameters showed a better reproducibility with mean %DIFF of −2% ± 5.4% for entropy and 1.8% ± 11.5% for homogeneity, where other textural parameters showed lower reproducibility ranging from 40.9% for lower limit to 62.7% for upper limit.[45,46]

Extraction of higher order textural features from NGTDM was investigated by Cook et al[47] using 48 baseline ^{18}F-FDG PET scans for patients with NSCLC treated with definitive chemoradiotherapy. The response was assessed 12 weeks after treatment on CT using the response evaluation criteria in solid tumours. Coarseness, contrast, busyness and complexity were extracted from NGTDM and compared with SUV measurements (mean, maximum and peak), metabolic tumour volume (MTV) and total lesional glycolysis (TLG) as predictors of response. The authors reported none of the SUV parameters nor TLG and MTV to be a significant predictor of responders and non-responders. Coarseness, contrast and busyness measurements were significantly different between responders and non-responders with an AUC of 0.80, 0.82 and 0.72, respectively. Nevertheless, the authors did not report whether any of the texture parameters could distinguish or predict PR and CR.[47]

Even though recent publications have shown a correlation between heterogeneity measurements and tissue response, the accuracy and precision of PET textural analysis is yet to be explored. Limitations owing to spatial resolution, noise, motion artefacts, image acquisition parameters and reconstruction methods are expected to lead to a degradation of the extracted textural features and should be addressed. A recent review on texture analysis in PET by Cook et al[48] focussing on technical factors and clinical application of radiomics in PET, supports this conclusion.

TEXTURE ANALYSIS IN MRI

Assessing tumour heterogeneity by utilizing dynamic contrast enhanced (DCE)-MRI in defining tumour functional risk volume (FRV) to predict treatment outcome in cervical cancer was investigated by Mayr et al[49] in 2012. FRV is defined as a region of low-contrast uptake, where the signal intensity is <2.1 when compared with that of the image before contrast injection. The authors presented a heterogeneity characterization approach that consisted generating relative signal intensity (RSI) distributions based on the perfusion levels of individual voxels that were then tabulated in an RSI histogram. The study included 102 patients diagnosed with cervical cancer stage IB_2-IVA based on the International Federation of Gynaecology and Obstetrics staging system of cervical cancer. The total FRV was derived from perfusion heterogeneity assessment prior to chemoradiotherapy and during treatment in Weeks 2–2.5 and 4–5. The FRV_2 was measured at 2–2.5 weeks during the chemoradiotherapy treatment course, while the FRV_3 was measured at 4–5 weeks during the chemoradiotherapy treatment course. FRVs were associated with disease-specific survival end points and primary tumour control. Patients' median follow-up was reported to be 6.8 years. The results reported by the authors suggest FRV to be a significant early predictor of outcome in the long term. Pre-treatment FRV was least predictive, where the power of prediction increased from 24.3% to 42.5% for FRV_2 and 45.2% for FRV_3 in discriminating 6-year actuarial tumour control rate and recurrences of primary tumour.[49,50] GLCM-based texture analysis in DCE-MRI[51] was reported to be successful in discriminating responders and non-responders for 89 patients diagnosed with locally advanced breast cancer and receiving neoadjuvant chemotherapy, its predictive value for patients undergoing radiotherapy is yet to be determined. Nevertheless, from 14 extracted features, only contrast, difference variance and difference entropy reached significance.[51]

CHALLENGES IN TEXTURE-BASED RADIOTHERAPY PLANNING

Whilst a customized radiotherapy plan based on prediction of tumour response to treatment is an attractive idea, texture analysis is hindered by multiple challenges. Uncertainties in patient positioning, organ motion, interobserver variability, image acquisition parameters, reconstruction algorithms and texture extraction methods affect the extracted textural features, and all must be investigated and addressed. Yet, the major question is the biological origin of image texture.

To date, the biological basis of textural analysis remains not fully understood. *In vivo* studies relating the apparent heterogeneity of tumours extracted from medical images to biological phenomena have not been attempted yet. Ganeshan et al[38] attempted to investigate an association of CT textural analysis with angiogenesis and hypoxia using pimonidazole staining and hypoxia marker glucose transporter protein (GLUT)-1 on surgically removed lesions. The analysis showed a correlation between the SD and the MPP in the ROI and pimonidazole staining. Uniformity of the distribution of positive grey-level pixel values showed an inverse association with GLUT-1, whereas MPP showed an inverse association with CD34 expression (this represents angiogenesis) in both contrast and non-contrast CT.[38] The study was based on measurements taken outside the human body on surgically removed lesions (*ex vivo*) and comprised a small cohort; hence a strong relationship could not be inferred.

Although the heterogeneous uptake of ^{18}F-FDG in tumours was reported to correlate with histopathology of untreated xenografted squamous cell carcinoma in a study conducted on mice,[52] the heterogeneous ^{18}F-FDG uptake can be caused by many factors as reported by van Velden et al.[53] Image noise and partial volume effect are major confounding factors in PET that lead to apparent heterogeneity in the tumour and necessitate correction before texture analysis can be performed. The authors recommended applying Van Cittert deconvolution, where the image is smoothed with a gaussian filter that represents the scanner point spread function, the resulting smoothed image is then subtracted from the original one and the ratio between the two images is added to a copy of the original image. The process is repeated in a loop until the filtered image is equivalent to the original image leading to a sharper PET image.[53–56] The suggested method showed significant improvement (Student's *t*-test, $p < 0.05$) on area under the normalized cumulative SUV–volume histogram curve. Yet, for the area under the non-normalized cumulative SUV-volume histogram, improvement did not reach significance. Furthermore, a recent study by Brooks and Grigsby[57] investigated the effect that a small size ROI has on the quantification of metabolic heterogeneity in PET. The authors investigated this hypothesis by creating a set of shapeless tumour test PET images where the intensity of the voxels were randomly assigned based on a known intensity distribution of PET images obtained from cervical tumour ^{18}F-FDG images. The test images set included 25 images per tumour volume. Local entropy was measured from a generated GLCM from each test image and then plotted against tumour volume. The authors reported an increase in entropy value as a function of tumour volume up to 45 cm^3 of the tumour volume. A flatness in entropy value was reported on tumour volume >45 cm^3. The authors concluded that a volume of 45 cm^3 (700 voxels) is needed to adequately sample the distribution of intensity and measure heterogeneity within a 95% confidence level.[57] The latter implications leave the reliability of ^{18}F-FDG-based texture analysis in doubt.

RESPIRATORY MOTION

Although a study by Vaidya et al[41] reported no significant difference in *p*-values in associating textural features extracted from motion-corrected PET and non-corrected PET images to tumour local failure, respiratory motion leads to image blurring, inaccurate representation of tumour volume and potentially mislocalization of lesions especially for tumours located in the thorax. A recent publication by Aristophanous et al[58] investigated target delineation in lung tumours on four-dimensional (4D) and 3D PET images. The authors reported the mean difference between tumour volume in 3D PET and tumour volume in 4D PET for lesions located in the lower lobe and lower mediastinum to be 50%, while the upper lobe difference was 10%. In addition, the lesions exhibiting motion >3 mm showed a larger difference (approximately 54%) than those lesions moving <3 mm (%DIFF 14%). The authors concluded that the benefit of 4D PET volume definition is dependent on tumour location and the range of motion.[58] While recently available PET/CT scanners have respiratory gating, there is not a currently established method for 4D MRI in radiotherapy. Many methods have been proposed in the literature for investigating the feasibility of 4D-MRI,[59–62] yet none has been clinically implemented. Overall, studies focusing on tumour motion are challenged by the lack of ground truth, where target delineation is subject to intra-observer variability. Furthermore, changes in patients' breathing patterns and irregular breathing are complicating factors, as well as patients' movement between consecutive scans. The magnitude of the effect of tumour motion on textural analysis is yet to be investigated since change in the apparent tumour volume would lead to a change in measured texture, especially when the extent of motion is similar in scale to studied textural features.

TUMOUR DELINEATION

Tumour delineation critically influences treatment outcome but whether it affects the power of extracted textural features is still under debate. According to Ng et al,[63] a large cross-sectional area of the tumour is sufficiently representative and provides comparable results to whole tumour analysis when applying histogram–filtration methods to predict survival. Yet, other studies in using texture analysis in tissue classification have reported volumetric texture analysis to be superior to 2D analysis owing to loss of spatial information.[64–66] Whether the precision of spatial information is needed in predicting survival and response is still to be explored further since most of the studies in the literature have adopted one methodology without comparison of other methods. On the other hand, it has been reported by Mattonen et al[43] that delineation of GGO in follow-up CT scans using expansion and contraction of 1–2 mm of the GGO had minimal effect on extracted textural features, leading to the assumption of lower impact of delineation variability in predicting recurrence.[43] Furthermore, a study by Leijenaar et al[67] in 2013 investigating the effect of interobserver variability on extracted textural features from ^{18}F-FDG PET reported 91% of the extracted textural features to have high stability using intra-class correlation coefficient (ICC). Medium stability was seen in 8% of the features, while low stability was reported in 1% of the features. The study included 23 patients with NSCLC, whose tumours were manually delineated by 5 independent observers on each fused PET-CT scan. The extracted textural features included first-order, second-order and higher order statistics summing to 44 textural features in total. The authors suggested ICC $\geq$ 0.8 to represent high stability, whereas 0.8 < ICC $\geq$ 0.5 is medium stability and ICC < 0.5 is low stability. The authors did not justify the threshold values used for high, medium and low ICC range and whether the variation has an effect on the utilization of these textural features as potential biomarkers.[67]

Nevertheless, intra-observer variability in target delineation should be minimized, and semi-automated and fully automated segmentation approaches have been suggested in the literature. Many of the research studies have been published investigating multiple PET and MRI segmentation methods, such as clustering, active contours and thresholding.[68] On the other hand, multimodality image segmentation methods were suggested by El Naqa et al[69] and Chowdhury et al.[70] The authors presented frameworks to link structures of interest across different imaging modalities to aid in target delineation for radiotherapy planning. Based on the published literature, the proposed methods suffer from shortcomings and rely heavily on image

quality, thus a recommendation cannot be easily reached. Yet, image segmentation is crucial in radiotherapy planning and should be implemented while considering the achievable accuracy level and the technical feasibility. Manual delineation of tumours is still considered the gold standard in radiotherapy planning.

Textural analysis is based on intensity values in the image, which is implicated by multiple technical factors owing to discrepancy between scanners, scanning protocols, reconstruction algorithm, noise and image post-processing, which will all affect extracted textural features. The robustness of texture analysis against the previously mentioned factors is yet to be explored.

FUTURE DIRECTION FOR TEXTURE ANALYSIS IN RADIOTHERAPY

Essentially, there is a need for more advanced quantitative image analysis methods in the field of radiotherapy. Initial results relating tumour texture to treatment outcome and patients' survival in radiotherapy are promising. Predicting tumour response prior to radiotherapy could potentially allow assigning of patients to high- and low-risk groups, where high-risk groups may be targeted with more aggressive treatment courses. In addition, predicting recurrence earlier from follow-up scans would be valuable for planning medical intervention as early as possible. Although a standardization of textural analysis in predicting survival is desirable, the methods explored in the literature did not explore all available possibilities to establish optimum standards in implementing texture analysis where confounding factors are addressed. Furthermore, the number of textural features that could be extracted from radiological images is numerous, yet, some were only investigated in the literature and even fewer showed an association or correlation with patient survival or response to treatment. The significance of these textural features should be investigated, and the casual relationship between these features and outcome prediction/survival should be understood.

Despite the fact that texture analysis based on intensity histograms reported promising results, it remains a global, spatially insensitive, simplistic approach for rather a complex question where global features are extracted. GLCM-based textural features reported the most significant results, yet quantizing GLCM to reduce its complexity leads to loss of spatial information. Two techniques that could be further explored are principal component analysis, which will reduce dimensionality and reduce computational cost, and wavelet transform, which iteratively decomposes the 2D image based on frequency and directionality of the signal into multiple components.[18] A combination of multi-approach texture analysis could be beneficial in overcoming the inherited shortcomings of available methods. Ideally, a certain standardization of texture analysis with regard to image quality and quantization methods is needed.

From the available literature, it is postulated that CT texture analysis will be the leading imaging modality in texture analysis given its superior resolution and its good integration into the field of radiotherapy compared with PET. Moreover, the poor resolution of PET, high noise, and tracer uptake and quantification challenges underpowers the extracted features and could lead to pseudotexture. MRI-based texture analysis seems an attractive idea owing to its superior soft tissue contrast, yet there are many challenging aspects to overcome. Target delineation based on fMRI is still not fully explored where published research studies investigating quantification of MRI functional parameters and target delineation are relatively scarce compared with that of CT imaging. In addition, the standardization of fMRI quantification is still not well established. fMRI is still not widely available and not as fully integrated into radiotherapy planning as CT and PET.

CONCLUSION

In conclusion, there is growing evidence of a correlation between tumour texture and treatment outcome providing the potential of tailored radiotherapy based on a tumour's predicted response. Medical images contain more information than is currently being utilized. Hence, texture analysis could enrich and complement existing methodologies, but this needs further validation. Texture analysis techniques should be developed aiming to quantify the heterogeneity of the tumours free from confounding factors. To date, only a few textural features showed significance in predicting outcome or patients' survival. However, the causal relationship is not fully understood and in need of further investigation.

REFERENCES

1. De Ruysscher D, Nestle U, Jeraj R, MacManus M. PET scans in radiotherapy planning of lung cancer. *Lung Cancer* 2012; **75**: 141–5. doi: 10.1016/j.lungcan.2011.07.018
2. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012; **366**: 883–92. doi: 10.1056/NEJMoa1113205
3. American Joint Committee on Cancer (AJCC). Cancer staging references. Available from: https://cancerstaging.org/references-tools/Pages/What-is-Cancer-Staging.aspx
4. Phelps ME. Positron emission tomography provides molecular imaging of biological processes. *Proc Natl Acad Sci U S A* 2000; **97**: 9226–33.
5. Brun E, Kjellén E, Tennvall J, Ohlsson T, Sandell A, Perfekt R, et al. FDG PET studies during treatment: prediction of therapy outcome in head and neck squamous cell carcinoma. *Head Neck* 2002; **24**: 127–35.
6. Hicks RJ, Mac Manus MP, Matthews JP, Hogg A, Binns D, Rischin D, et al. Early FDG-PET imaging after radical radiotherapy for non-small-cell lung cancer: inflammatory changes in normal tissues correlate with tumor response and do not confound therapeutic response evaluation. *Int J Radiat Oncol Biol Phys* 2004; **60**: 412–18.
7. Xue F, Lin LL, Dehdashti F, Miller TR, Siegel BA, Grigsby PW. F-18 fluorodeoxyglucose uptake in primary cervical cancer as an indicator of prognosis after radiation therapy.

Gynecol Oncol 2006; **101**: 147–51. doi: 10.1016/j.ygyno.2005.10.005

8. Bollineni VR, Widder J, Pruim J, Langendijk JA, Wiegman EM. Residual ^{18}F-FDG-PET uptake 12 weeks after stereotactic ablative radiotherapy for stage I non-small-cell lung cancer predicts local control. *Int J Radiat Oncol Biol Phys* 2012; **83**: e551–5.
9. Oh D, Lee JE, Huh SJ, Park W, Nam H, Choi JY, et al. Prognostic significance of tumor response as assessed by sequential ^{18}F-fluorodeoxyglucose-positron emission tomography/computed tomography during concurrent chemoradiation therapy for cervical cancer. *Int J Radiat Oncol Biol Phys* 2013; **87**: 549–54.
10. Higgins KA, Hoang JK, Roach MC, Chino J, Yoo DS, Turkington TG, et al. Analysis of pretreatment FDG-PET SUV parameters in head-and-neck cancer: tumor SUVmean has superior prognostic value. *Int J Radiat Oncol Biol Phys* 2012; **82**: 548–53. doi: 10.1016/j.ijrobp.2010.11.050
11. Kidd EA, Thomas M, Siegel BA, Dehdashti F, Grigsby PW. Changes in cervical cancer FDG uptake during chemoradiation and association with response. *Int J Radiat Oncol Biol Phys* 2013; **85**: 116–22. doi: 10.1016/j.ijrobp.2012.02.056
12. van Loon J, Offermann C, Ollers M, van Elmpt W, Vegt E, Rahmy A, et al. Early CT and FDG-metabolic tumour volume changes show a significant correlation with survival in stage I–III small cell lung cancer: a hypothesis generating study. *Radiother Oncol* 2011; **99**: 172–5.
13. Petit SF, Aerts HJ, van Loon JG, Offermann C, Houben R, Winkens B, Ollers MC, et al. Metabolic control probability in tumour subvolumes or how to guide tumour dose redistribution in non-small cell lung cancer (NSCLC): an exploratory clinical study. *Radiother Oncol* 2009; **91**: 393–8. doi: 10.1016/j.radonc.2009.02.020
14. Hicks RJ, MacManus MP, Matthews JP, Hogg A, Binns D, Rischin D, et al. Early FDG-PET imaging after radical radiotherapy for non-small-cell lung cancer: inflammatory changes in normal tissues correlate with tumor response and do not confound therapeutic response evaluation. *Int J Radiat Oncol Biol Phys* 2004; **60**: 412–18.
15. Hicks RJ. Role of ^{18}F-FDG PET in assessment of response in non-small cell lung cancer. *J Nucl Med* 2009; **50**(Suppl. 1): 31–42S. doi: 10.2967/jnumed.108.057216
16. Kalff V, Duong C, Drummond EG, Matthews JP, Hicks RJ. Findings on ^{18}F-FDG PET scans after neoadjuvant chemoradiation provides prognostic stratification in patients with locally advanced rectal carcinoma subsequently treated by radical surgery. *J Nucl Med* 2006; **47**: 14–22.
17. 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: 10.1016/j.ejca.2011.11.036
18. Materka A, Strzelecki M. *Texture analysis methods—a review. COST B11 report*. Brussels, Belgium: Technical University of Lodz, Institute of Electronics; 1998.
19. O'Connor JP, Rose CJ, Jackson A, Watson Y, Cheung S, Maders F, et al. DCE-MRI biomarkers of tumour heterogeneity predict CRC liver metastasis shrinkage following bevacizumab and FOLFOX-6. *Br J Cancer* 2011; **105**: 139–45. doi: 10.1038/bjc.2011.191
20. Hayano K, Lee SH, Yoshida H, Zhu AX, Sahani DV. Fractal analysis of CT perfusion images for evaluation of antiangiogenic treatment and survival in hepatocellular carcinoma. *Acad Radiol* 2014; **21**: 654–60. doi: 10.1016/j.acra.2014.01.020
21. Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, Cook GJ. Quantifying tumour heterogeneity in ^{18}F-FDG PET/CT imaging by texture analysis. *Eur J Nucl Med Mol Imaging* 2013; **40**: 133.
22. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern* 1973; **3**: 610–21.
23. Galloway MM. Texture analysis using grey level run lengths. *Comput Graph Imaging* 1974; **4**: 172–9.
24. Amadasun M, King R. Textural features corresponding to textural properties. *IEEE Trans Syst Man Cybern* 1989; **19**: 1264–74.
25. 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
26. Vasconcelos V, Silva JS, Marques L, Barroso J. Statistical textural features for classification of lung emphysema in CT images: a comparative study. Proceedings of the 5th Iberian Conference on Information Systems and Technologies (CISTI); 16–19 June 2010; Santiago de Compostela, Spain. Piscataway, NJ; IEEE, 2010. pp. 1–5.
27. Xu Y, Sonka M, Mclennan G, Guo J, Hoffman EA. MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. *IEEE Trans Med Imaging* 2006; **25**: 464–75. doi: 10.1109/TMI.2006.870889
28. Gibbs P, Turnbull LW. Textural analysis of contrast-enhanced MR images of the breast. *Magn Reson Med* 2003; **50**: 92–8. doi: 10.1002/mrm.10496
29. Yu H, Caldwell C, Mah K, Mozeg D. Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. *IEEE Trans Med Imaging* 2009; **28**: 374–83.
30. Eliat P, Olivié D, Saïkali S, Carsin B, Saint-Jalmes H, de Certaines JD. Can dynamic contrast-enhanced magnetic resonance imaging combined with texture analysis differentiate malignant glioneuronal tumors from other glioblastoma? *Neurol Res Int* 2012; **2012**: 195176.
31. Ganeshan B, Miles KA, Young RC, Chatwin CR. Hepatic enhancement in colorectal cancer: texture analysis correlates with hepatic hemodynamics and patient survival. *Acad Radiol* 2007; **14**: 1520–30. doi: 10.1016/j.acra.2007.06.028
32. Ganeshan B, Miles KA, Young RC, Chatwin CR. Hepatic entropy and uniformity: additional parameters that can potentially increase the effectiveness of contrast enhancement during abdominal CT. *Clin Radiol* 2007; **62**: 761–8.
33. Miles KA, Ganeshan B, Griffiths MR, Young RC, Chatwin CR. Colorectal cancer: texture analysis of portal phase hepatic CT images as a potential marker of survival. *Radiology* 2009; **250**: 444–52. doi: 10.1148/radiol.2502071879
34. Ganeshan B, Miles KA, Young RC, Chatwin CR. Texture analysis in non-contrast enhanced CT: impact of malignancy on texture in apparently disease-free areas of the liver. *Eur J Radiol* 2009; **70**: 101–10. doi: 10.1016/j.ejrad.2007.12.005
35. Ganeshan B, Abaleke S, Young RC, Chatwin CR, Miles KA. Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage. *Cancer imaging* 2010; **10**: 137–43.
36. Ganeshan B, Panayiotou E, Burnand K, Dizdarevic S, Miles K. Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. *Eur Radiol* 2012; **22**: 796–802. doi: 10.1007/s00330-011-2319-8
37. Ganeshan B, Skogen K, Pressney I, Coutroubis D, Miles K. Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival. *Clin Radiol* 2012; **67**: 157–64. doi: 10.1016/j.crad.2011.08.012
38. 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: 10.1148/radiol.12112428
39. Yip C, Landau D, Kozarski R, Ganeshan B, Thomas R, Michaelidou A, et al. Primary

esophageal cancer: heterogeneity as potential prognostic biomarker in patients treated with definitive chemotherapy and radiation therapy. *Radiology* 2014; **270**: 141–8.

40. Win T, Miles KA, Janes SM, Ganeshan B, Shastry M, Endozo R, et al. Tumor heterogeneity and permeability as measured on the CT component of PET/CT predict survival in patients with non-small cell lung cancer. *Clin Cancer Res* 2013; **19**: 3591–9. doi: 10.1158/1078-0432.CCR-12-1307
41. Vaidya M, Creach KM, Frye J, Dehdashti F, Bradley JD, El Naqa I. Combined PET/CT image characteristics for radiotherapy tumor response in lung cancer. *Radiother Oncol* 2012; **102**: 239–45. doi: 10.1016/j.radonc.2011.10.014
42. Mattonen SA, Palma DA, Haasbeek CJ, Senan S, Ward AD. Distinguishing radiation fibrosis from tumour recurrence after stereotactic ablative radiotherapy (SABR) for lung cancer: a quantitative analysis of CT density changes. *Acta Oncol* 2013; **52**: 910–18.
43. Mattonen SA, Palma DA, Haasbeek CJ, Senan S, Ward AD. Early prediction of tumor recurrence based on CT texture changes after stereotactic ablative radiotherapy (SABR) for lung cancer. *Med Phys* 2014; **41**: 033502. doi: 10.1118/1.4866219
44. El Naqa I, Grigsby P, Apte A, Kidd E, Donnelly E, Khullar D, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern Recognit* 2009; **42**: 1162–71. doi: 10.1016/j.patcog.2008.08.011
45. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges JP, et al. Intratumor heterogeneity characterized by textural features on baseline ^{18}F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med* 2011; **52**: 369–78.
46. 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 ^{18}F-FDG PET. *J Nucl Med* 2012; **53**: 693–700. doi: 10.2967/jnumed.111.099127
47. Cook GJ, Yip C, Siddique M, Goh V, Chicklore S, Roy A, et al. Are pretreatment ^{18}F-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: 10.2967/jnumed.112.107375
48. Cook GJR, Siddique M, Taylor BP, Yip C, Chicklore S, Goh V. Radiomics in PET: principles and applications. *Clin Transl Imaging* 2014; **2**: 269–276. doi 10.1007/s40336-014-0064-0.
49. Mayr NA, Huang Z, Wang JZ, Lo SS, Fan JM, Grecula JC, et al. Characterizing tumor heterogeneity with functional imaging and quantifying high-risk tumor volume for early prediction of treatment outcome: cervical cancer as a model. *Int J Radiat Oncol Biol Phys* 2012; **83**: 972–9.
50. Mayr NA, Yuh WT, Arnholt JC, Ehrhardt JC, Sorosky JI, Magnotta VA, et al. Pixel analysis of MR perfusion imaging in predicting radiation therapy outcome in cervical cancer. *J Magn Reson Imaging* 2000; **12**: 1027–33.
51. Ahmed A, Gibbs P, Pickles M, Turnbull L. Texture analysis in assessment and prediction of chemotherapy response in breast cancer. *J Magn Reson Imaging* 2013; **38**: 89–101. doi: 10.1002/jmri.23971
52. Henriksson E, Kjellen E, Wahlberg P, Ohlsson T, Wennerberg J, Brun E. 2-deoxy-2-F-^{18}fluoro-D-glucose uptake and correlation to intratumoral heterogeneity. *Anticancer Res* 2007; **27**: 2155–9.
53. 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: 10.1007/s00259-011-1845-6
54. Teo BK, Seo Y, Bacharach SL, Carrasquillo JA, Libutti SK, Shukla H, et al. Partial-volume correction in PET: validation of an iterative postreconstruction method with phantom and patient data. *J Nucl Med* 2007; **48**: 802–10. doi: 10.2967/jnumed.106.035576
55. Tohka J, Reilhac A. A Monte Carlo study of deconvolution algorithms for partial volume correction in quantitative PET. *IEEE Nucl Sci Symp Conf* 2006; **6**: 3339–45.
56. Geets X, Lee JA, Bol A, Lonneux M, Grégoire V. A gradient-based method for segmenting FDG-PET images: methodology and validation. *Eur J Nucl Med Mol Imaging* 2007; **34**: 1427–38. doi: 10.1007/s00259-006-0363-4
57. Brooks FJ, Grigsby PW. The effect of small tumor volumes on studies of intratumoral heterogeneity of tracer uptake. *J Nucl Med* 2014; **55**: 37–42. doi: 10.2967/jnumed.112.116715
58. Aristophanous M, Berbeco RI, Killoran JH, Yap JT, Sher DJ, Allen AM, et al. Clinical utility of 4D FDG-PET/CT scans in radiation treatment planning. *Int J Radiat Oncol Biol Phys* 2012; **82**: e99–105. doi: 10.1016/j.ijrobp.2010.12.060
59. Cai J, Chang Z, Wang Z, Paul Segars W, Yin FF. Four-dimensional magnetic resonance imaging (4D-MRI) using image-based respiratory surrogate: a feasibility study. *Med Phys* 2011; **38**: 6384–94. doi: 10.1118/1.3658737
60. Tryggestad E, Flammang A, Han-Oh S, Hales R, Herman J, McNutt T, et al. Respiration-based sorting of dynamic MRI to derive representative 4D-MRI for radiotherapy planning. *Med Phys* 2013; **40**: 051909. doi: 10.1118/1.4800808
61. Tokuda J, Morikawa S, Haque HA, Tsukamoto T, Matsumiya K, Liao H, et al. Adaptive 4D MR imaging using navigator-based respiratory signal for MRI-guided therapy. *Magn Reson Med* 2008; **59**: 1051–61. doi: 10.1002/mrm.21436
62. Blackall JM, Ahmad S, Miquel ME, McClelland JR, Landau DB, Hawkes DJ. MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. *Phys Med Biol* 2006; **51**: 4147–69. doi: 10.1088/0031-9155/51/17/003
63. 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: 10.1016/j.ejrad.2012.10.023
64. Mahmoud-Ghoneim D, Toussaint G, Constans JM, de Certaines JD. Three dimensional texture analysis in MRI: a preliminary evaluation in gliomas. *Magn Reson Imaging* 2003; **21**: 983–7.
65. Xu Y, Sonka M, McLennan G, Guo J, Hoffman E. Sensitivity and specificity of 3-D texture analysis of lung parenchyma is better than 2-D for discrimination of lung pathology in Stage 0 COPD. *Proc SPIE Med Imaging*; 2005; **5746**: 474–485.
66. Chen W, Giger ML, Li H, Bick U, Newstead GM. Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images. *Magn Reson Med* 2007; **58**: 562–71. doi: 10.1002/mrm.21347
67. 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: 10.3109/0284186X.2013.812798
68. Klein S, van der Heide UA, Lips IM, van Vulpen M, Staring M, Pluim JP. Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Med Phys* 2008; **35**: 1407–17.
69. El Naqa I, Yang D, Apte A, Khullar D, Mutic S, Zheng J, et al. Concurrent multimodality image segmentation by active contours for radiotherapy treatment planning. *Med Phys* 2007; **34**: 4738–49.
70. Chowdhury N, Toth R, Chappelow J, Kim S, Motwani S, Punekar S, et al. Concurrent segmentation of the prostate on MRI and CT via linked statistical shape models for radiotherapy planning. *Med Phys* 2012; **39**: 2214–28. doi: 10.1118/1.3696376