ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

Annals of Oncology 28: 1191–1206, 2017
doi:10.1093/annonc/mdx034
Published online 7 February 2017

REVIEW

Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology

E. J. Limkin[1,2†], R. Sun[1,2,3†], L. Dercle[4], E. I. Zacharaki[5], C. Robert[1,2,3], S. Reuzé[1,2,3], A. Schernberg[1,2,3], N. Paragios[5,6], E. Deutsch[1,2] & C. Ferté[1,7*]

[1]Radiomics team, INSERM U1030, Gustave Roussy; [2]Department of Radiotherapy, Gustave Roussy, Paris-Saclay University, Villejuif; [3]Faculty of Medicine, Paris Sud University, Kremlin-Bicetre; [4]Department of Nuclear Medicine and Endocrine Oncology, Gustave Roussy, Paris-Saclay University, Villejuif; [5]Center for Visual Computing, CentraleSupelec/Paris-Saclay University/Inria, Châtenay-Malabry; [6]TheraPanacea, Paris; [7]Department of Head and Neck Oncology, Gustave Roussy, Paris-Saclay University, Villejuif, France

Correspondence to: Dr Charles Ferté, Radiomics Team, INSERM U1030, Gustave Roussy, Paris-Saclay University, Villejuif, France. Tel: +33-1-42-11-46-17; E-mail: charles.ferte@gustaveroussy.fr

†Both authors contributed equally as senior authors.

Medical image processing and analysis (also known as Radiomics) is a rapidly growing discipline that maps digital medical images into quantitative data, with the end goal of generating imaging biomarkers as decision support tools for clinical practice. The use of imaging data from routine clinical work-up has tremendous potential in improving cancer care by heightening understanding of tumor biology and aiding in the implementation of precision medicine. As a noninvasive method of assessing the tumor and its microenvironment in their entirety, radiomics allows the evaluation and monitoring of tumor characteristics such as temporal and spatial heterogeneity. One can observe a rapid increase in the number of computational medical imaging publications—milestones that have highlighted the utility of imaging biomarkers in oncology. Nevertheless, the use of radiomics as clinical biomarkers still necessitates amelioration and standardization in order to achieve routine clinical adoption. This Review addresses the critical issues to ensure the proper development of radiomics as a biomarker and facilitate its implementation in clinical practice.

Key words: computational medical imaging, radiomics, quantitative imaging, precision medicine, tumor biology

Introduction

Medical images possess valuable information which can be harnessed through computer assisted interpretation. This technique, termed radiomics, is a rapidly-emerging discipline with the goal of extracting quantitative data from medical images to be used as clinical decision support tools [1–3]. In the context of oncology, information obtained from standard imaging modalities [computed tomography scan (CT), magnetic resonance imaging (MRI), and Positron emission tomography scan (PET)], usually refers to simple traits such as gross shape, contrast enhancement, and size. However, imaging information is much richer, and the goal of radiomics is to extract high throughput quantitative features, covering the fields of texture, advanced shape modeling, and heterogeneity, to name a few. The increasing resolution quality has led to three-dimensional (3D) image acquisitions containing millions of voxels available for analysis, making the development of radiomics a natural progression, as more data necessitated increased computing capabilities to harness more information.

Radiomics has immense potential to improve knowledge in tumor biology and guide the management of patients at bedside [4]. Medical image analysis allows tumor monitoring across time, with images being routinely acquired throughout the course of treatment. Thus, imaging biomarkers may be used for and contribute to cancer detection, diagnosis, choice of therapeutic strategy, prognosis inference, prediction of response, and surveillance. Tumors exhibit spatial heterogeneity and temporal variation, recognized as major causes of treatment failure and modulators of intrinsic tumor aggressiveness [5–7]. Imaging

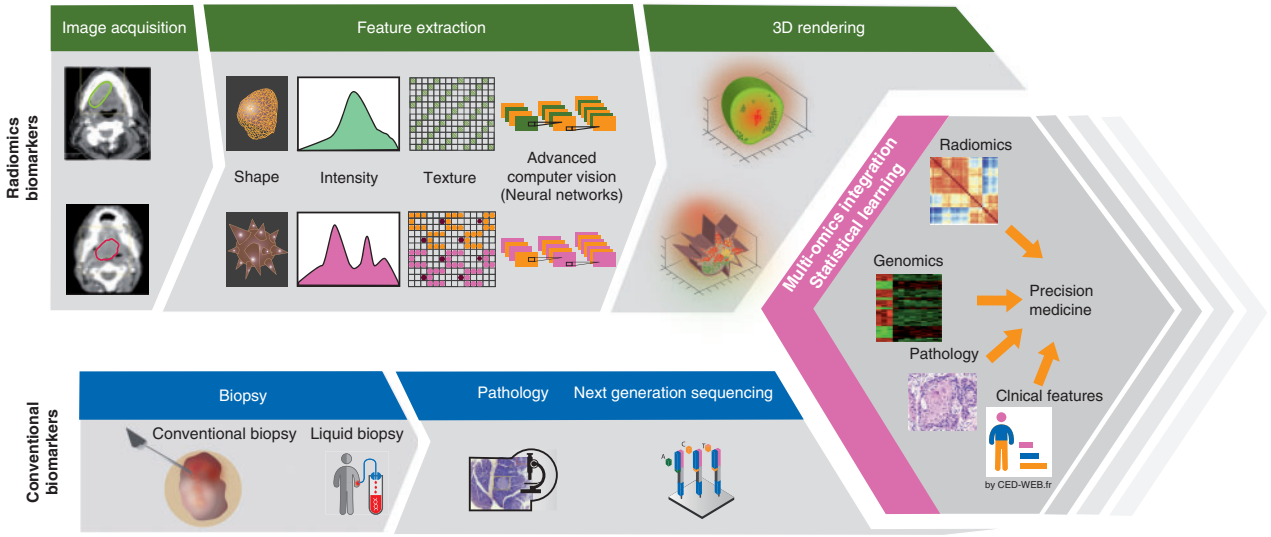

Figure 1. The radiomics pipeline, showing the major steps: image acquisition and segmentation, feature extraction, three-dimensional (3D) rendering and statistical learning. Important to note is that radiomics data is meant to be integrated and analyzed with clinical, pathologic, and –omics data to improve precision medicine.

allows assessment of the entire tumor plus surrounding tissue—it is not blind to global heterogeneity—as opposed to invasive needle biopsies that are limited by sampling site. Radiomics could serve as a 'virtual biopsy' that would provide complementary information to, but not replace, conventional biopsies which remain vital in deep genomic analysis.

This review will systematically go through the radiomics pipeline, the significant milestones achieved in oncology, and future perspectives for improvement. The aim is to increase the awareness and interest of the oncology community to radiomics, demystify the field for non-imaging experts, and engage the community to be involved in its further development and adoption in clinical practice. We will demonstrate the potential of radiomics to advance precision medicine.

Radiomics pipeline

The process of radiomics consists of discrete steps: image acquisition and segmentation, feature extraction, statistical learning and 3D rendering (Figure 1).

Image acquisition and segmentation

Radiomics may be applied to different/multiple modalities, and selecting the one(s) to investigate depends on several factors. Radiomics quantitatively explores the distribution of signal intensities within a region or volume of interest (ROI/VOI). The spatial resolution of images varies, being around 1 mm for CT and MRI and 4 mm for PET. ROIs that are too small (e.g. sub-centimeter nodules) may not provide sufficient voxel information for analysis, whereas ROIs that are too large may be impacted by tumor heterogeneity (e.g. large tumors often present hypoxia in their centers). Standardization and calibration of non-ionizing procedures (Ultrasound, MRI) are intrinsically more complex than techniques based on photon detection (PET or CT) [8]. The minimal concentration of a molecule that is detectable on PET is 10^{-12} moles compared with 10^{-3} moles for MRI, meaning that the former might be 10^{9} more sensitive for *in vivo* molecular imaging. Patient motion and respiration during acquisition affects the quality of reconstructed images; a CT is usually acquired in a few seconds whereas functional (MRI) or molecular (PET) imaging lasts several minutes. Data might be extracted from retrospective standard-of-care images, leading to large pool of patients. However, acquisition parameters vary considerably, which can introduce signal variations not due to biologic effects [8–10]. As with other high throughput technologies such as genomics, the aggregation of multiple datasets in radiomics can lead to substantial unwanted effects on the data. In determining the inclusion criteria for studies, an option is to have a large cohort in order to be less sensitive to variations due to acquisition/reconstruction parameters, or have a small cohort with homogenous data.

Crucial in the process is correct delineation of ROIs to be analyzed. Segmentations must be reproducible and reliable. Automatic methods are preferable for precision and efficiency. It has been proven that inter- and intra-observer variabilities are lower in automatic versus semi-automatic versus manual delineation [10–13]. However, semi-automatic delineation is usually mandatory, since automated methods are feasible only if there are strong signal differences between the lesion and the background. This is observed in PET and CT of pulmonary tumors, and possibly in certain MRI sequences. Commonly, such as in tumors surrounded by relatively homogenous normal structures, an experienced physician is required to correct contours, entailing computer-aided outline detection followed by manual correction.

Currently, there are several open source platforms equipped with automatic and semi-automatic contouring functions, such as 3DSlicer (Growcut algorithm) [14], which have active online support, and are continuously updated. Generated contours should be stored in an easily utilizable format for analysis across various platforms. Commonly used is DICOM-RTSTRUCT,

which enables data sharing within and outside of the radiotherapy workflow, containing information on the images and ROIs [15]. Other formats such as analyze [16] and NIfTI are also used [17].

Feature extraction

Preprocessing. Raw imaging data need pre-processing to discriminate the signal from the noise. One optional step is filtering the signal within the ROI, which defines the frequencies to be utilized for subsequent analysis [18]. The choice of filter is guided by the nature of the imaging modality and tumor tissue. Another is the discretization or resampling [19, 20] of signal intensities that partitions continuous voxel values to finite/nominal intervals called bins. Techniques involve either absolute (using a fixed bin size) or relative (using a fixed number of bins whose size depends on the minimum and maximum values within the tumor) discretization. The choice of method is crucial as extracted features vary accordingly [21, 22]. Several studies [19, 20, 23] have shown that absolute discretization results to features with better repeatability and lower sensitivity to changes, with the added advantage of not being volume dependent.

Radiomics features. Features may be classified into several categories. There are quantitatively extracted descriptors of size, shape, and other radiologic terminologies which characterize the tumor surface. First-order statistics are used to study the distribution of voxel values without considering spatial relationships [24]; second-order statistics characterize spatial relationships between voxels, initially described by Haralick [25] such as the co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM) [26], gray-level size zone matrix (GLZLM) [27], and the neighborhood gray-level different matrix (NGLDM) [28]. Filter grids such as Gabor and Fourier may be used both in the pre-processing step and for extracting spatial or spatio-temporal features [29, 30]. A limitation is that some extracted values are dependent on the ROIs contoured.

The extracted features can be global (one value for the whole ROI), or local (a value per image patch) when inhomogeneous patterns are present in the image, where dimensionality significantly increases if simple concatenation of local descriptors is carried out. For this, more advanced frameworks explore compact statistical representations based on coding structures/dictionaries. When visual vocabularies and visual word weights are learned jointly, performance can be improved, as shown in classification of breast tissue density in mammograms, lung tissue in high-resolution CT, and brain tissue in MRI [31]. A more detailed review on texture analysis methods focusing on microscopy images of cells or tissues can be found in [32].

The stability and the accuracy of features should be confirmed through the use of test–retest datasets; a good practice policy is to eliminate features that prove to be unreliable in the test–retest. To this end, several datasets are publicly available. Of note is the RIDER [33] dataset, which allows validation of results in the same set of patients with two scans taken 15 min apart.

Statistical learning

The impact of the high number of radiomics variables. The current radiomics pipeline typically incorporates around 50–5000 quantitative features (p), and these are still expected to increase. Meanwhile, the number of patients (n) in studies remains small, leading to a situation where $p \gg n$, or the 'curse-of-dimensionality' [34]; resulting in a high probability of false positive results [35]. Adjustments for multiple comparisons (Bonferroni correction [36]) and controlling the false discovery rate (Benjamini–Hochberg [37]) are commonly utilized methods to address this. Another issue is overfitting, which can be reduced by cross-validation with independent training and validation datasets [34]. Several techniques of dimensionality reduction can be used to reduce the number of variables for analysis by exploiting statistical correlations and data redundancy [38]. Unsupervised techniques map the data through a linear (e.g. prinicipal and independent component analyses) or non-linear (e.g. ISOMAP, locally linear embedding) transformation in a lower-dimensional space, such that information loss is minimized, whereas supervised techniques select a subset of the original variables such that prediction accuracy is maximized. Feature selection can be carried out independently, before classification or regression, or be combined into a single mathematical problem (e.g. Lasso, ElasticNet) [39]. Attention has to be given in the case of supervised learning and small datasets to not overfit the data; feature selection should be carried out externally to the cross-validation procedure to correctly estimate the empirical error.

Promising machine learning approaches for prediction and classification tasks. Machine learning approaches [40] such as decision trees and random forests [41–46], support vector machines [39, 47, 48] and more recently deep neural networks [49] appear to be promising in the domain of medical image computing. The recent increase of available annotated imaging data in public portals expedited the use of these techniques. In particular, deep convolutional neural networks (CNNs) have been shown to excel at learning a hierarchy of increasingly complex features directly from raw data, alleviating the explicit extraction of low-order, pre-defined features. In such frameworks, feature extraction and selection are carried out jointly with classification within the optimization of the same deep architecture, thus performance can be tuned in a systematic fashion. Common CNN schemes train patch-level classifiers [50] that automatically locate discriminative regions and then aggregate local predictions. The learning process is usually facilitated by pre-training using standardized data followed by supervised training for fine-tuning.

Thus far, CNNs have been shown to excel in the detection and classification of pulmonary nodules in large series of lung CTs from the Lung Image Database Consortium (LIDC-IDRI) [41, 43, 51]. A comparison in mortality prediction from chest CTs [52] between (i) a unified deep learning framework (features and classifier are automatically learned in a single optimization process) and (ii) a standard multi-stage framework (pre-defined radiomics features are introduced into a classifier), showed increased accuracy of the deep learning framework by 2.5%–12.5%. Similar concepts have been applied to mammograms for breast cancer screening [53, 54].

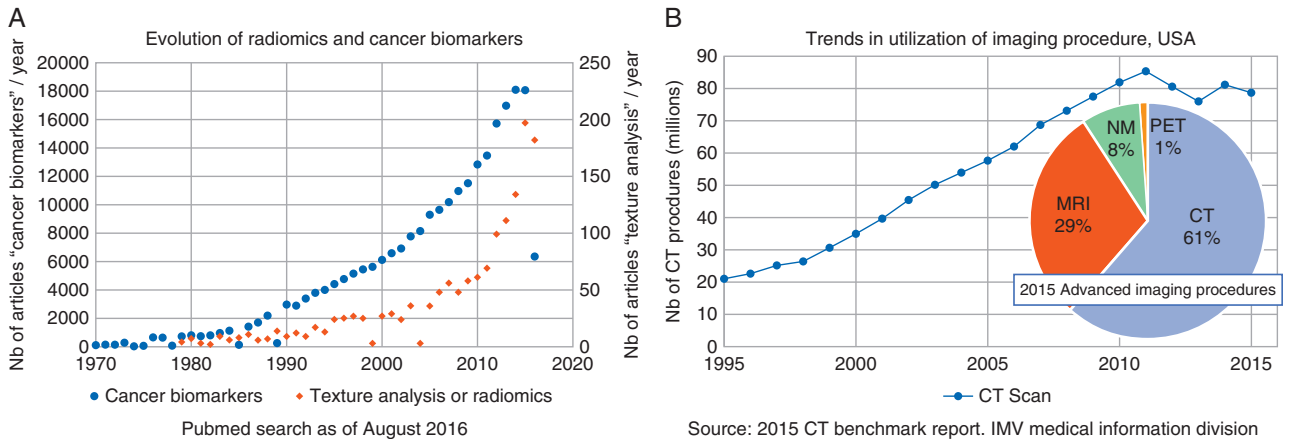

Figure 2. (A) Comparison of publications on radiomics and molecular biomarkers. (B) Trends in imaging utilization in the USA. CT, Computed tomography scan; MRI, Magnetic resonance imaging; NM, Nuclear medicine; PET, Positron emission tomography.

Three-dimensional rendering

The visualization of the entire tumor, commonly carried out in diagnostic radiology, is possible through 3D rendering [55]. Radiomics takes this a step further through visualization and assessment of tumor heterogeneity which provides invaluable information in clinical oncology and in cancer research. For instance, this can direct which intra-tumoral region is best to biopsy for adequate samples. In radiotherapy dose painting, rendering is instrumental in visualizing treatment-resistant regions, facilitating dose escalation and consequently decreasing normal tissue toxicity [56]. Theragnostic imaging [57], the use of molecular/functional imaging to prescribe radiation distribution, will benefit from radiomics features reflecting phenotypes associated with poor radio-responsiveness [58].

Significant milestones of radiomics in oncology

Radiomics is substantial in oncology, not surprising given the increased availability of and information in patient imaging data. Oncologic imaging is a substantial part of a radiologist's daily workload. More than 128 million imaging examinations were realized in the United States in 2015, greater than 60% of which were CTs (Figure 2B); and an increase of >36% in use of advanced imaging was seen from 2005 to 2015 [59], reiterating the vast amount of available data. This section aims to depict the impact of radiomics in each stage of cancer care, with tables outlining key details.

Advancements in the understanding of tumor biology

Tumor heterogeneity. The extensive genetic and molecular landscape within tumors—tumor heterogeneity—is known to be specific to the malignant process. Intra-tumoral heterogeneity is dynamic and is modified by therapeutic effects. Heterogeneity may portend treatment resistance and poor outcomes due to the emergence of resistant subclonal populations [60]. Through quantitative serial analysis of imaging, both temporal and spatial heterogeneities may be analyzed. In particular, texture analysis is emerging as an effective method to quantify heterogeneity. Table 1 displays some representative work. To illustrate, in CTs of lung adenocarcinomas, entropies of the tumor core and boundary regions were computed separately, with results showing that a higher ratio between the two are associated with poorer outcome [61]. Thus, this imaging phenotype may reflect distinct traits such as necrosis in the core and proliferation in the periphery. In PET, texture analysis showed that healthy lung tissues are more homogenous than malignancies, and texture features differentiated tumor histologies [20]. Similar studies on texture heterogeneity, with or without other feature classes, depict quantitative imaging traits mirroring genomic and molecular phenotypes [62–69].

Modeling of key oncogenic processes. Several biological processes are known to be necessary components of oncogenesis [70, 71]. Various molecular biology techniques have been developed to better appraise biological and genetic causes of these, and the emergence of quantitative imaging analysis is a promising tool to complement and enhance existing techniques.

A study in lung cancer used an angiogenesis-related marker injected in specimens, and linear correlations were shown between CT texture heterogeneity features and percentage of the tumor stained for the marker [72]. In glioblastomas, the 'contrast enhancement imaging phenotype' was significantly associated with tumor angiogenesis [64]. In breast cancers, certain computer extracted imaging phenotypes of MRIs have been shown to differentiate subtypes [62].

In lung cancers, multi-categorical quantitative imaging features were shown to reflect tumor hypoxia [63, 66]. In the preclinical setting, causal relationship between genetic changes and imaging features was demonstrated by pre-imaging administration of doxycycline which induces hypoxic changes [73]. Clinically, these results may be helpful in identifying tumor resistance and the need for intensified regimens.

Table 1. Advancements in the prediction of tumor phenotypes, genotypes, molecular subtypes through radiomics

Tumor type and imaging modality	Author, year	Training/ validation sets (*N*)	Extracted radiomics features	Feature selection and statistical learning	Biologic correlation and relevance
Breast cancer (TCGA-TCIA dataset) CE MRI	Li, 2016 [62]	T: 91	38 CEIPs (size, shape, morphology, enhancement texture, kinetic curve, enhancement-variance kinetics)	Leave-one-case-out with stepwise feature selection and linear discriminant analysis, logistic regression	CEIPs are able to predict breast cancer molecular subtypes (ER, PR, HER2Neu, basal-like) with AUCs from 0.67 to 0.89.
Breast Cancer (TCGA –TCIA dataset) MRI	Zhu, 2015 [74]	T: 91	38 radiomics features (size, shape, morphology, enhancement texture, kinetic curve, enhancement-variance kinetics)	Linear regression, Benjamini–Hochberg adjustment, affinity propagation clustering	Landmark description of the associations between radiomics features and somatic mutations, CNVs, gene expression and miRNA. Radiomic phenotypes correlated with key oncogenic pathways such as DNA replication.
Breast cancer (TCGA –TCIA dataset) CE MRI	Mazurowski, 2014 [65]	T: 48	23 initial features: 4 geometric, 14 Haralick texture, 1 kinetic and 4 dynamic	Logistic regression analysis, Bonferroni correction	Radiomics features have significant association with the luminal B breast cancer molecular subtype ($P = 0.0015$)
NSCLC PET	Orlhac, 2015 [20]	T: 48 V: phantoms	7 Textural indices (homogeneity, entropy, SRE, LRE, RLNU, LGLZE, HGLZE); SUV max and metabolic volume; 3 texture matrices (co-occurrence, gray-level run length and gray-level zone length matrices)	Spearman correlations, no adjustment for multiple testing	Texture features differentiate between benign and malignant tissues, and between different cancer histological subtypes. The presence of necrosis was significantly different between squamous and adenocarcinomas, unlike the T or N stage.
NSCLC CT	Ganeshan, 2013 [63]	T: 14	Low-order texture features computed upon the pixel intensity distribution. Laplacian of Gaussian transformation, spatial scale filters.	Linear mixed-effects model approach, Holm- adjustment	Texture features (standard deviation and mean value of positive pixels) are associated with tumor hypoxia [Glut1/pimonidazole] and angiogenesis [CD34]
NSCLC CT PET	Gevaert, 2012 [66]	T: 26	153 computational image features, 26 semantic image features, and a PET SUV	Generalized linear regression with lasso regularization	Image features associate with publicly available gene expression pathways (Hypoxia, KRAS pathway) with accuracies of 59%–83%
Colorectal cancer xenografts CE CT	Panth, 2015, (pre-clinical) [73]	33 mice	625 radiomics features (intensity, shape, texture, wavelet features, Laplacian of Gaussian features)	Intra-class correlation coefficiant rank, based on the 50% top-ranked features	Radiomics features demonstrate the causality between genomics changes and imaging features
Glioblastoma multiforme CE MRI	Jamshidi, 2014 [67]	T: 23	(1) Infiltrative versus edematous T2 abnormality, (2) degree of contrast enhancement, (3) necrosis, (4) supraventricular zone (SVZ) involvement, (5) mass effect, and (6) contrast-to-necrosis ratio.	Resampling statistics, analysis of variance, Pearson correlation coefficient	Gene-to-trait associations were found such as contrast-to-necrosis ratio with KLK3 and RUNX3; SVZ involvement with the Ras oncogene family and the metabolic enzyme TYMS; and vasogenic edema with the oncogene FOXP1 and PIK3IP1

Continued

Table 1. *Continued*

Tumor type and imaging modality	Author, year	Training/ validation sets (*N*)	Extracted radiomics features	Feature selection and statistical learning	Biologic correlation and relevance
Glioblastoma Multiforme CE MRI	Diehn, 2008 [64]	T: 22 V: 110	10 binary imaging traits (enhancement, necrosis, mass effect,T2 edema, cortical involvement, SVZ involvement, C:N ratio, contrast/T2 ratio,T2 edema, T2 heterogeneity)	Unsupervised hierarchical clustering; Spearman rank-correlation coefficient	Significant associations were found between: angiogenesis and tumor hypoxia with the contrast enhancement imaging phenotype ($P = 0.012$); proliferation gene-expression signature and mass effect phenotype ($P = 0.0017$); EGFR protein overexpression and contrast:necrosis imaging trait ($P < 0.002$)
Glioblastoma Multiforme CE MRI	Zinn, 2011 [75]	T: 26 V: 26	Quantitative models of edema/ invasion, enhancing tumor, necrosis	For imaging-genomic analysis: Comparative Marker Selection, Ingenuity Pathway Analysis	Imaging traits associated with upregulation of mRNA involved in cellular migration/invasion (PERIOSTIN), which was seen to correlate with decreased survival ($P = 0.0008$)
Soft tissue sarcoma CT	Hayano, 2015 [68]	T: 20	Low-order texture features computed upon the pixel intensity distribution. Laplacian of Gaussian transformation, spatial scale filters	Multi-variate Cox regression model; no correction for multiple testing	Tumor texture features are associated with microvessel density, VEGF, soluble VEGF receptor-1, and overall survival, and the mean value of positive pixels is an independent prognostic factor in multi-variate analysis ($P = 0.01$)
Renal Cell Cancer (RCC) (TCGA–TCIA dataset) CE CT	Zhu, 2016 [69]	T: 112	PADUA scoring system: Exophytic, Longitudinal, Rim location, Renal sinus, UCS, Tumor size, Face location	Stratified according to PADUA score; correlation of the PADUA system and its radiological features with miRNA molecular subtypes GSEA, Pearson metric	Imaging features of PADUA scores may serve as molecular surrogates for RCC diagnosis, prognosis and personalized treatment of patients with specific genomic profiles. Higher PADUA scores were significantly associated with epithelial to mesenchymal transition pathways ($P < 0.001$)

CT, computed tomography scan; MRI, magnetic resonance imaging; PET, Positron emission tomography; CE, contrast-enhanced; TCGA-TCIA, The Cancer Genome Atlas-The Cancer Imaging Archive; CEIPs, computer-extracted image phenotypes; ER, estrogen receptor; PR, progesterone receptor; CNVs, copy number variations; NSCLC, non-small cell lung cancer; SRE, short-run emphasis; LRE, long-run emphasis; RLNU, run length non-uniformity; LGLZE, low gray-level zone emphasis; HGLZE, high gray-level zone emphasis; SUV, standardized uptake value; EGFR, epidermal growth factor receptor; GSEA, gene set enrichment analysis; AUC, area under the curve; VEGF, vascular endothelial growth factor.

Further, radiomics phenotypes have been shown to correlate with mRNA and protein expression in breast MRIs [74], alluding to the underlying mechanisms of tumor invasion. In glioblastomas, peritumoral fluid attenuation inversion recovery MRI signal abnormalities were found to reflect genes and microRNAs accounting for cellular migration and invasion [75]. In addition, lymphocyte infiltration has been reflected in imaging features of HER2+ breast cancers [76].

Radiomics features also appear to model tumor proliferation. A landmark study in liver cancer demonstrated that the combination of 28 traits could reconstruct as much as 78% of transcriptome variation, and were specifically associated with genes

involving cell proliferation [77]. In breast cancers, analysis of TCGA-TCIA data revealed clear associations between radiomics phenotypes and proliferation at the protein and gene expression levels. Imaging phenotypes of increased tumor size were also associated with P-cadherin expression, shown to correlate with increased proliferation [74].

Better implementation of precision medicine

Precision medicine is an approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle; integrating information from multiple sources with the end goal of personalized management [78]. In this section, some key radiomics applications are described.

Diagnosis. One of the biggest challenges in oncology is the development of accurate, cost–effective screening procedures. In the National Lung Cancer Screening Trial (NLST), low-dose CTs interpreted by radiologists have been shown to be beneficial in mortality reduction in a cohort of 53 000 patients [79]. These however exhibit high false-positive rates (>90%); wherein radiomics could be effective in improving specificity. Using NLST data, patients with screening-detected lung cancer were matched with subjects with benign nodules. Accuracies of 80% and 79% were found for predicting nodules that will become cancerous at 1 and 2 years, respectively [42]. Using LIDC-IDRI patients with a total of 42 340 lesions (through data augmentation), deep convolutional sequencers detected malignant nodules with diagnostic accuracy, sensitivity, and specificity >75% [41]. Two studies [43, 80] with 914 and 1375 lesions, respectively, published promising results with predictive performances ranging from 60% to 80%. Notably, multi-scale CNNs attained 86.84% accuracy on classification with automatic nodule detection and segmentation [43]. Other publications in lung carcinomas reveal similar findings [46, 81, 82]. Table 2 presents representative publications concerning diagnosis.

Staging and prognosis. Accurate staging determines the aggressiveness of therapeutic strategy and spells the difference between curative and palliative treatments, requiring constant advancements in imaging techniques to improve decision-making. In colorectal carcinomas, for which liver metastases are frequent, radiomics features extracted from unenhanced hepatic CT scans showed texture abnormalities suggestive of metastases in apparently disease-free areas [83]. This could streamline the staging process and minimize treatment delays with the information steering the clinician toward doing confirmatory examinations for patients at risk. Plus, early detection of metastases may increase chances of complete tumor eradication.

In prognostication, increased shape complexity in lung adenocarcinoma [61] was associated with poorer clinical outcomes. Morphologically similar tumors by visual inspection turned out to have large differences in quantitative parameters, denoting that radiomics can supplement radiologists' interpretations. Another study [84] demonstrated that radiomics features are prognostic for both distant metastasis and survival, and a radiomics signature significantly improves prognostication when added to clinical data. Similar studies were published in colorectal cancers [85, 86]. In a large cohort involving multiple datasets with varying tumor types [4], differences in imaging phenotypes showed clinical significance and impact on prognosis. These results were reiterated in a dataset of almost 900 patients [87] and externally validated in oropharyngeal cancers [88]. Table 3 shows recent radiomics publications on tumor staging and prognosis.

Prediction of treatment response. Radiomics features could be used to predict which patients would respond to a treatment regimen. Table 4 summarizes notable publications. In glioblastomas, a robust association between radio-phenotypes and gene expression has been shown, including a link with epidermal growth factor receptor (EGFR) overexpression [64, 89]. Radiomics features have been shown to stratify treatment outcomes from angiogenic therapy in recurrent glioblastomas [90]. EGFR mutation [91] and response to Gefinitib [92] were also reflected in a combination of features in lung cancers. With the rapid rise in targeted therapy, it is worthwhile to continue discovering radiogenomic associations that may influence management. An imaging surrogate to could aid patient selection and avoid unwarranted expense and toxicities for non-responders.

In the neoadjuvant chemotherapy setting, a multi-parametric model in breast MRIs identified non-responders with 84% sensitivity [93], with the goal of developing a computer-assisted prediction solution, which may be more cost–effective than molecular assessments. In rectal carcinomas, a combination of radiomics features was seen to be predictive of pathologic response to neoadjuvant treatment [49]. Studies in other cancer localizations [94–97] have shown the possibility of assessing treatment response using imaging markers.

In the advent of immunotherapy, there have been been patients who experience pseudoprogression (PSPD) [98], which has been shown to be due to lymphocyte infiltration in and around the tumor [99]. A study demonstrated that a radiomics signature from CTs could reflect tumors with increased lymphocytes and discriminate PSPD from true progression [100], possibly aiding decision-making for equivocal cases.

Disease monitoring and surveillance. Inflammation leads to post-treatment reactions that might complicate response evaluation by imaging. In this context, functional imaging helps differentiate scar tissue from viable tumor, but equivocal cases remain. Radiomics can further enhance evaluation. For instance, texture features from CT images of lung cancer treated with stereotactic ablative body radiotherapy (SABR) showed that the ground glass appearance (an area of hazy increased lung opacity through which vessels and bronchial structures may still be seen [101]), following SABR predicts recurrence versus radiation-induced lung injury, which has a similar radiologic picture [102]. This is particularly useful because lung cancer typically progresses quickly; hence the decision for salvage therapy is most valuable if made early, likely providing more treatment options compared with late-detected disease. In addition, a pilot study on renal cell cancers has demonstrated the possibility of capturing sub-visual treatment-related imaging changes [103].

Table 2. Advancements in oncologic diagnosis in radiomics

Tumor type and imaging modality	Author, year	Training/validation sets (*N*)	Extracted radiomics features	Feature selection and statistical learning	Clinical utility for diagnosis
Lung cancer Low dose CT	Liu, 2016 [46]	T: 102 V: 70	24 semantic radiological traits (location, size, shape, margin, density, internal features, external features, associated findings) used to create a linear classification model	Accuracy/AUROC, Hold-out cross-validation, Bootstrap, Youden J index	Radiological image traits are useful in predicting malignancy in lung nodules (accuracies ranging from 70% to 77%)
Lung cancer (NLST dataset) CT	Hawkins, 2016 [42]	malignant: 104(T) + 92(V); benign: 208(T) + 196 (V)	219 three-dimensional features (size, shape, location, and texture matrices)	Hierarchical learning framework—multi-scale convolutional neural networks, random forest classifier	Radiomics features are effective in the prediction of malignant versus benign lung nodules with accuracies of 80% and 79%.
Lung cancer CE CT	Wu W., 2016 [81]	T: 198 V: 152	440 radiomic features (voxel intensity distribution, shape, texture matrices)	Correlation-based feature elimination and univariate feature selection. 24 feature selection and 3 classification methods tested	Radiomics features predict NSCLC histology (highest achieved AUC = 0.72)
Lung cancer (LIDC-IDRI dataset) CT	Kumar, 2015 [41]	93 patients, 42 340 lesions T: 34 295 V: 3810	abstract imaging-based features learned from a deep convolutional neural network learning architecture	Deep convolutional neural network; binary decision tree classifier	Method used outperforms the state-of-the art approach for lung nodule benign versus malignant classification (accuracy 77%, sensitivity 79%, specificity 76%)
Lung cancer (LIDC-IDRI dataset) CT	Shen, 2015 [43]	T: 1100 nodules V: 275 nodules (with augmentation)	Learned classifiers from the neural network algorithm	Multi-scale convolutional neural networks	The convolutional neural network method used showed 86% accuracy for pulmonary nodule classification
Lung adenocarcinoma High resolution CT	Maldonado, 2013 [82]	T: 54 nodules V: 86 nodules	Computer-aided nodule assessment and risk yield (CANARY): representative exemplars of the spectrum of solid and ground glass components of nodules	Affinity propagation (unsupervised clustering), Multinomial logistic regression, nonparametric Spearman correlation	CANARY can noninvasively characterize pulmonary nodules of the adeno-carcinoma spectrum, with the exemplar distribution within each nodule correlating well with the proportion of histologic tissue invasion ($P < 0.0001$)
Lung cancer (LIDC-IDRI dataset) CT	Zinovev, 2011 [80]	914 instances T: 90% V: 10%	63 2D low-level image features from four categories: shape, texture, intensity, size	Multiple-label belief decision trees, fivefold cross validation	Multiple-label classification algorithms are an appropriate method of representing the diagnoses of radiologists on lung CT scans (AUC 69%)
Prostate cancer MRI	Fehr, 2015 [120]	T: 147	First- and second-order texture statistics from intensity distribution(mean, SD, skewness,kurtosis); Haralick features (energy, entropy, correlation, homogeneity, contrast)	Oversampling approach, support vector machine	Texture features predicts benign versus malignant prostate lesions and Gleason score of malignant tumors, with a better accuracy (93%) than using diffusion MRI alone

CT, computed tomography scan; MRI, magnetic resonance imaging; CE, contrast-enhanced; AUROC, area under the receiver-operating characteristic; NLST, national lung cancer screening trial; LIDC-IDRI, Lung Image Database Consortium; SD, standard deviation; AUC, area under the curve.

Table 3. Recent radiomics works in staging and prognosis

Tumor type and imaging modality	Author, year	Training/ validation sets (N)	Extracted radiomics features	Feature selection and statistical learning	Utility and significance
NSCLC and HNSCC CE CT	Aerts, 2014 [4]	T: 31, 21, 422 V: 225, 136, 95, 89	440 radiomics image features (intensity, shape,texture, multi-scale wavelet)	Selection of the most stable variables ($n = 100$), unsupervised clustering; Friedman test, bootstrap approach	Significant associations were found between imaging features and tumor stage and overall survival, with the radiomics signature having a CI = 0.65 (NSCLC), CI = 0.69 (HNSCC)
NSCLC and HNSCC CE CT	Parmar, 2015 [87]	T: Lung - 422 HN - 136 V: Lung 2 - 225 HN2 - 95	440 radiomics features (tumor intensity, shape, texture, wavelet)	Consensus clustering	Significant agreement in the clusters between training and validation sets was seen. Features were associated with stage (lung AUC = 0.61, H&N AUC = 0.77), HPV status (H&N AUC = 0.58) and prognosis (lung AUC = 0.60, H&N AUC = 0.68).
Lung adenocarcinoma CT	Coroller, 2015 [84]	T: 98 V: 84	635 radiomics features (intensity, shape, texture, Laplacian of Gaussian and Wavelet filtered)	Bootstrapping, Benjamini–Hochberg correction	Imaging features aid in the identification of patients at risk of developing distant metastases (CI = 0.61), facilitating individual treatment decisions
Lung adenocarcinoma CE CT	Grove, 2015 [61]	T: cohort 1 = 61; cohort 2 = 47 V: 32 (RIDER test–retest)	shape complexity (convexity) and intratumoral density variation (entropy ratio)	Concordance correlation coefficient, Dynamic Range, correlation matrix	Increasing tumor entropy and lower convexity are associated with overall survival, even after adjustment for tumor stage.
Glioblastoma Multiforme CE MRI	Kickingereder, 2016 [90]	T: 112 V: 60	4842 total 17 first-order features, 9 volume and shape features, 162 texture features	Supervised principal component analysis, Cox proportional hazard models, Integrated Brier scores	Radiomics-based classification of recurrent glioblastoma permits the prediction of treatment outcome to anti-angiogenic therapy through PFS ($P = 0.030$) and OS ($P = 0.001$).
Colorectal cancer CE CT	Huang, 2016 [85]	T: 326 V: 200	150 texture features, 24 feature-based radiomics signature; Radiomics nomogram	LASSO; Multi-variate binary logistic regression, nomograms and calibration plots	Radiomics nomogram predicts lymph node metastases (CI = 0.78), beneficial in pre-treatment decisions.
Colorectal cancer CE CT	Liang, 2016 [86]	T 286 V: 208	150 texture features; 16-feature-based radiomics signature	LASSO; the radiomics score	Texture features can be utilized in preoperative staging of colorectal carcinomas (AUC = 0.71).
Oropharyngeal Squamous cell Cancer (OPSCC) CT	Leijenaar, 2015 [88]	V: 542	radiomics signature (energy, Compactness, Gray level non-uniformity, Wavelet)	Cox proportional hazards model	Results show the applicability as a prognostic index of a radiomics signature which was trained in lung and H&N and cancer validated well in an external cohort of OPSCC (CI = 0.63)
Breast Cancer (TCGA -TCIA dataset) CE MRI	Li,2016 [121]	T: 84	38 CEIPs (size, shape, morphology, enhancement texture, kinetic curve, enhancement-variance kinetics)	Leave one-case-out cross-validation analysis with logistic regression	CEIPs predict the recurrence risk as assessed by Oncotype Dx, PAM50 or Mammaprint (AUCs from 0.55 to 0.88)

CT, computed tomography scan; MRI, magnetic resonance imaging, CE, contrast-enhanced; TCGA-TCIA, The Cancer Genome Atlas-The Cancer Imaging Archive; NSCLC, non-small cell lung cancer; HNSCC, head and neck squamous cell carcinoma; CI, concordance index; AUC, area under the curve; LASSO, least absolute shrinkage and selection operator; PFS, progression-free survival; OS, overall survival; CEIPs, computer-extracted image phenotypes.

Table 4. Publications on the prediction of treatment response through radiomics

Tumor type and imaging modality	Author, year	Training/ validation sets (*N*)	Extracted radiomics features	Feature selection and statistical learning	Utility and significance
Lung adenocarcinomas High resolution CT	Aerts, 2016 [92]	T: 47 V: 31 (Rider test–retest)	183 initial features, 11 independent features used (volume, Gabor Energy, Sigmoid Function, Shape Index, Boundary Radius, GLCM, Laws Energy)	Spearman rank statistic, AUC, intraclass correlation coefficient	Radiomics features predict EGFR mutation status and associated response to Gefitinib at baseline (AUC = 0.67) and in change in pre- and post treatment (AUC = 0.74–0.91) CT scans of NSCLC
Lung adenocarcinomas CE CT	Y Liu, 2016 [91]	T: 385	Semantic radiologic features (location, size, shape, margin, attenuation, internal, enhancement, external, associated findings) assessed by 3 radiologists	Multiple logistic regression analyses, backward elimination method, AUROC	CT features of lung adenocarcinomas can be an image biomarker for EGFR mutation status, with the use of clinical variables combined with CT features (AUROC = 0.778) being superior to use of clinical variables alone (AUROC = 0.690)
NSCLC PET	Cook, 2015 [95]	T: 47	First-order and high-order primary tumor texture features	Cox and logistic regression analyses	Reduced quantitative heterogeneity features (percentage change in entropy) in PET scans are associated with time to disease progression ($P = 0.03$) and treatment response ($P = 0.01$)
NSCLC CE CT	Mattonen, 2015 [102]	T: 22	Mean density, first-order texture, energy, entropy, correlation, inverse difference moment, inertia, cluster shade, and cluster prominence	Independent samples *t*-test with unequal variances, Kolmogorov–Smirnov test, Wilcoxon signed rank test linear Bayes normal classifier, Spearman rank correlation coefficients	First and second-order texture features can predict eventual cancer recurrence based on CT images acquired within 5 months of SABR treatment (accuracies of 73%–77%)
Glioblastoma Multiforme (TCGA -TCIA dataset) CE MRI	Lee, 2015 [89]	T: 65	36 spatial habitat diversity (regions with distinctly different intensity characteristics) features based on pixel abundances w/in ROIs	Overall coefficient of variation; symbolic regression method	Features had association with overall survival (AUC = 0.74) and EGFR + (AUC = 0.85) status and could be a useful prognostic tool for MRIs of patients with gloiblastomas
Breast cancer CE MRI	Michoux, 2015 [93]	T: 69	20 texture, 3 kinetic, BI-RADS and biologic parameters	Logistic regression model, k-means clustering algorithm based on a nearest-cluster approach; leave-one-out cross validation	Radiomics features predicts response to neoadjuvant chemotherapy (accuracy = 68%), may be used in treatment decision-making
Breast cancer CE MRI	Ahmed, 2012 [94]	T: 100	Texture features based on co-occurrence matrices (Haralick features, cluster shade, cluster prominence)	Mann–Whitney, *t*-tests Categorized according to their chemotherapeutic response and histopathology	Certain texture parameters are significantly associated with treatment response (best-performing features $P = 0.039–0.048$) and tumor histology (best-performing features $P = 0.001–0.012$)
Rectal cancer CE MRI	Nie, 2016 [49]	T: 48	103 imaging features (texture, shape, histogram)	Mean-value based and voxelized analysis techniques; three-layer perceptron artificial neural network, feed-forward-back-propagation learning	Features reflect response to neoadjuvant therapy (AUC = 0.71–0.79 in voxelized analysis) and could influence treatment plans

Continued

Table 4. *Continued*

Tumor type and imaging modality	Author, year	Training/ validation sets (*N*)	Extracted radiomics features	Feature selection and statistical learning	Utility and significance
Rectal cancer PET	Bundschuh, 2014 [97]	T: 27	First-order primary tumor texture features (coefficient of variation, skewness, kurtosis)	ROC analysis, Youden index	Texture features reflect response to neoadjuvant chemoradiotherapy and prognostic capability for disease progression (AUC = 0.89 for the coefficient of variation feature)
Renal Cell Cancer CT	Goh, 2011 [96]	T: 39	First-order statistics (Entropy and uniformity), TEXRAD	Cox regression model	Heterogeneity biomarkers are associated with treatment response and time to progression (AUC = 0.71); with prediction rates better than standard response criteria (RECIST, Choi)
Renal Cell Cancer (RCC) integrated PET/MRI	Antunes, 2016 [103]	2 test/retest scans + mid treatment scan	66 radiomic features (raw T2w signal, post-processed T2w, 30 post-processed T2w textures, raw ADC map, 30 ADC textures, SUV, 2 PET textures)	Cox proportional hazards model	SUV and both T2w and ADC texture features appear to be able to capture subvisual TKI treatment-related changes in RCCs, with the highest-ranked radiomics feature yielding a normalized percentage change of 63% within the RCC region

CT, computed tomography scan; MRI, magnetic resonance imaging; PET, Positron emission tomography; CE, contrast-enhanced; TCGA-TCIA, The Cancer Genome Atlas-The Cancer Imaging Archive; NSCLC, non-small cell lung cancer; GLCM, gray level co-concurrence matrix; SABR, stereotactic ablative body radiotherapy; EGFR, epidermal growth factor receptor; AUC, area under the curve; T2, Weighted MRI; ADC, apparent diffusion coefficient; AUROC, area under the receiver operating characteristic.

Perspectives related to the effective translation of the radiomics biomarkers into the clinic

There is convincing evidence that radiomics could be an invaluable tool in revolutionizing oncology. Significant progress has been made, but further improvements are imperative to achieve routine utilization from bench to bedside.

Standardization and the perspective relative to molecular biomarkers

Radiomics literature is constantly growing, and we predict that it will follow the curve of the molecular biomarkers as interests and funding increase (Figure 2A). At present, however, the existing level of evidence is insufficient (Figure 3). There are notable differences in terms of sample size, methodology, performance metrics, and clinical utility; reiterating that improvements are essential.

As investigators have learned from discovery of biomarkers [104, 105], there are pitfalls to be avoided. In the same way as the *REporting recommendations for tumour MARKer prognostic studies* (REMARK) [106] or the *Minimum information about a microarray experiment* (MIAME) [107] guidelines, recommendations specific to radiomics are necessary. In reporting, key elements should be sufficiently detailed and made available to allow comparisons and validation: (i) raw imaging data including acquisition parameters, (ii) ROIs, (iii) radiomics features and the extraction software, methods, formulae used, and (iv) statistical learning methods. Meriting attention is the non-standardized names of extracted features, such that two publications might discuss a feature with the same formula/definition but call these differently.

The recently published roadmap for imaging biomarkers [108] is a notable advancement, showcasing key recommendations for clinical translation of radiomics. Also admirable are large initiatives aiming to develop automatic segmentation solutions such as the Google-National Health Service partnership *DeepMind Health* project [109]. Another is the *Grand Challenges in Biomedical Image Analysis* [110], with goals of developing algorithms for specific problems such as the *Lung Nodule Analysis (LUNA) Challenge* [111], a large-scale automatic nodule detection with 888 patients and *The Digital Mammography (DREAM) Challenge* [112] aiming to improve predictive accuracy of digital mammography with over 640 000 images.

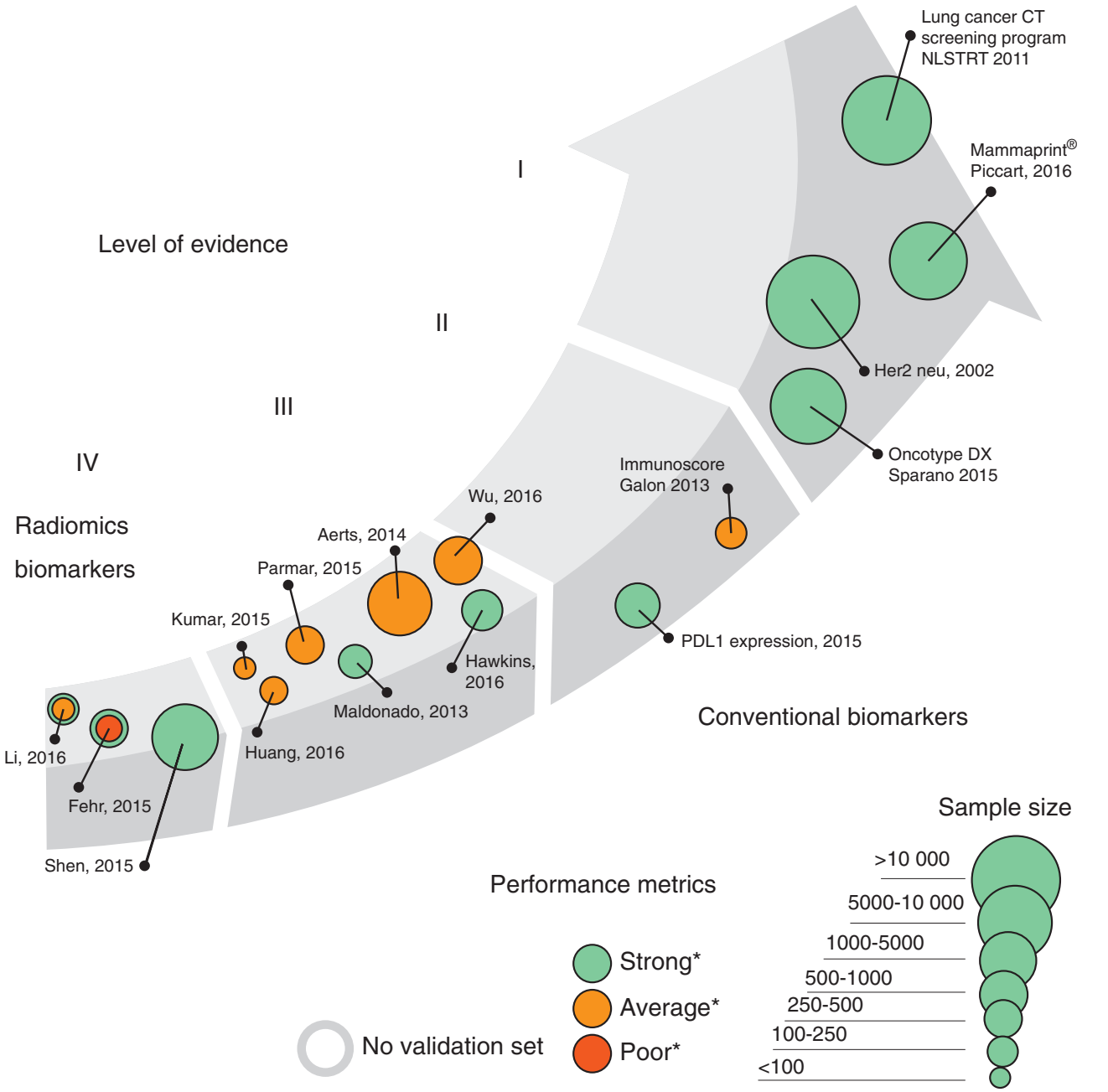

Figure 3. Comparison of key biomarker and radiomics studies. Shown are representative studies in the field of molecular biomarkers and imaging biomarkers, illustrating that the areas radiomics can be improved on in the future are (i) use of prospective cohorts (ii) increased number of subjects (iii) use of validation sets. Performance Metrics (Concordance indices/ Area under the curve): Strong >80, Average: 60–80, Poor <60.

To ensure robustness and dissemination of radiomics-based predictive tools, standardization of imaging protocols and feature calculations is ideal, but seldom attainable [10, 113]. Imaging carried out at different centers leads to bias [10, 114]. In fact, significant variation of radiomics features was observed across different CT scans through a phantom experiment [113]. The use of credentialing scanners, careful study design, noise reduction, and statistical analyses adjusted to account for unwanted effects are possible solutions. Table 5 highlights the current challenges and corresponding recommendations in radiomics.

Actual data access

Improving data management is challenging for a number of reasons including (i) administrative (manpower), (ii) ethical (patient privacy), and (iii) personal/institutional (intellectual

Table 5. Issues and solutions for radiomics studies

Step	Pitfall	Solution
Image acquisition and reconstruction	Differences in acquisition parameters	Need for standardized protocols adapted to each modality (CT, PET, MRI); and comprehensive description of the parameters being used.
		Exclusion of images with outlier acquisition parameters.
	Contrast enhancement protocols vary across machines and across patients (sarcopenia, adipose level, heart rate, etc.).	Standardized control ROIs such as muscle
Image segmentation	Intra/inter-observer variability	Semi-automatic segmentation with human correction/improvement
	Time consuming contouring methods	Development of (semi-)automated contouring methods (ideally open source)
Feature extraction	Large range of voxel intensities and image noise	Filtering procedure aiming to preserve the signal and reduce the unwanted noise
	Different discretization methods producing different results	Using fixed bin sizes (absolute discretization)
	Volume dependence	Testing for correlation between radiomics variables and volume
Statistical learning	Large number of features, small population ($p \gg n$) resulting into a high probability of false positives results and overfitting, the 'curse of dimensionality'	Bonferroni, Benjamini–Hochberg corrections
		Cross-validation
		Dimensionality reduction through supervised and unsupervised (PCA[a], ICA[b], ISOMAP, LLE[c]) techniques
	Feature selection and classification uncertainties, susceptibility to human error	Advanced machine learning approaches such as neural networks
General	Reproducibility is limited	Publications should include access to raw data, segmented ROIs, methods used for feature extraction
		A repository should be initiated containing imaging data, radiomics features, extraction software, methods, formulae and statistical learning methods

[a]Prinicipal component analysis.
[b]Independent component analyses.
[c]Locally linear embedding.

property). One successful undertaking is TCIA [115], through which investigators have access to robust anonymized imaging data in easily utilizable DICOM format.

Researchers should be encouraged to submit data to a centralized online radiomics repository akin to the Gene Expression Omnibus [116] for microarrays. A standardized non-software dependent format of storing and annotating data will facilitate multi-platform utilization. These will not only be useful before starting research projects, but also during and after; ensuring the integrity and availability of information. These must allow incorporation of image features, annotations, medical information, and genetic data in order to create prognostic and predictive models correlating imaging with genetic phenotypes and clinical outcomes. It is critical that the fidelity of the data is maintained and access is regulated, for which competent system administrators are mandatory. These are big steps, but we believe they are essential for the advancement of the discipline. Forming a multi-national consortium may be a prudent solution, some of whose functions would be to (i) draft guidelines on data collection, anonymization, and sharing, (ii) standardize reporting of radiomics studies.

Improving multi-disciplinary network and dissemination of radiomics

Significant efforts have been made to address these issues. The National Cancer Institute, in cooperation with other societies like the Canadian Institute of Health Research, Cancer Research United Kingdom and American College of Radiology Imaging Network, have supported initiatives, including the Quantitative Imaging Network (QIN), to promote the development of QI methods, annotated image databases, and QI standards [117]. Other significant efforts include the Quantitative Imaging Biomarkers Alliance (QIBA), a critical component of US' Cancer Moonshot initiative [118], and the Euregional Computer Assisted Theragnostics project (EuroCAT) [119], whose goals include enhancing data sharing and facilitating patient recruitment in clinical trials. More initiatives are necessary, with multi-disciplinary working groups that include oncologists, radiologists, medical physicists, applied mathematicians, and computer scientists, to improve the field and educate people on its use such that it can become a reliable part of a decision support system in oncology. Radiomics has been gaining ground in terms of exposure and interest in recent scientific congresses, with the number of publications per year almost doubling in the last three and almost tripling in the last five years (197 in 2015, an increase of 77%

since 2013 and 186% since 2011) (Figure 2A). This is translated in increased exposure of radiomics in current radiology meetings (e.g. RSNA) and in major oncology meetings (e.g. ASCO, ESMO, AACR, ASTRO). This is a formidable start but efforts need to be increased.

Discussion

Conclusion

Imaging biomarkers constructed from quantitative image analysis have great potential to advance precision medicine and to enhance cancer biology knowledge. As radiomics cements its position in translational cancer research to attain utilization at bedside, we anticipate radiomics data being integrated and analyzed with genomics, proteomics and other -omics; providing information invaluable in personalized medicine. Radiomics will certainly progress further with the advent of more imaging data, better algorithms, and availability of other data types such as coherent datasets integrating imaging, clinical, and genomic information.

Acknowledgement

The authors wish to thank C. Verjat for preparation of the figures.

Funding

The authors of this review received no grant from any funding agency; no grant number is applicable.

Disclosure

The authors have declared no conflicts of interest.

References

1. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016; 278(2): 563–577.
2. Kumar V, Gu Y, Basu S et al. QIN "Radiomics: The Process and the Challenges." Magn Reson Imaging 2012; 30(9): 1234–1248.
3. Lambin P, Rios-Velazquez E, Leijenaar R et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012; 48(4): 441–446.
4. Aerts HJWL, Velazquez ER, Leijenaar RTH et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014; doi:10.1038/ncomms5006.
5. Heppner GH. Tumor heterogeneity. Cancer Res 1984; 44(6): 2259–2265.
6. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. Nature 2013; 501(7467): 346–354.
7. O'Connor JPB, Rose CJ, Waterton JC et al. Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome. Clin Cancer Res 2015; 21(2): 249–257.
8. Bushberg JT, Boone JM. The Essential Physics of Medical Imaging. Philadelphia: Lippincott Williams & Wilkins 2011.
9. Galavis PE, Hollensen C, Jallow N et al. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. Acta Oncol 2010; 49(7): 1012–1016.
10. Zhao B, Tan Y, Tsai W-Y et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. Sci Rep 2016; 6: 23428.
11. Balagurunathan Y, Kumar V, Gu Y et al. Test–retest reproducibility analysis of lung CT image features. J Digit Imaging 2014; 27(6): 805–823.
12. Parmar C, Rios Velazquez E, Leijenaar R et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. PLoS One 2014; doi:10.1371/journal.pone.0102107.
13. Rios Velazquez E, Aerts HJWL, Gu Y et al. A semiautomatic CT-based ensemble segmentation of lung tumors: comparison with oncologists' delineations and with the surgical specimen. Radiother Oncol 2012; 105(2): 167–173.
14. Velazquez ER, Parmar C, Jermoumi M et al. Volumetric CT-based segmentation of NSCLC using 3D-Slicer. Sci Rep 2013; doi:10.1038/srep03529.
15. Gorthi S, Bach Cuadra M, Thiran J-P. Exporting contours to DICOM-RT structure set. Insight J 2009.
16. FormatAnalyze - MRC CBU Imaging Wiki. http://imaging.mrc-cbu.cam.ac.uk/imaging/FormatAnalyze (7 February 2017, date last accessed).
17. NIfTI-1 Data Format — Neuroimaging Informatics Technology Initiative. http://nifti.nimh.nih.gov/nifti-1 (7 February 2017, date last accessed).
18. Wells WM. Efficient synthesis of Gaussian filters by cascaded uniform filters. IEEE Trans Pattern Anal Mach Intell 1986; PAMI-8(2): 234–239.
19. Leijenaar RTH, Nalbantov G, Carvalho S et al. The effect of SUV discretization in quantitative FDG-PET radiomics: the need for standardized methodology in tumor texture analysis. Sci Rep 2015; doi:10.1038/srep11075.
20. Orlhac F, Soussan M, Chouahnia K et al. 18F-FDG PET-derived textural indices reflect tissue-specific uptake pattern in non-small cell lung cancer. PLoS ONE 2015; 10(12): e0145063.
21. Brooks FJ, Grigsby PW. The effect of small tumor volumes on studies of intratumoral heterogeneity of tracer uptake. J Nucl Med. 2014; 55(1): 37–42.
22. Orlhac F, Soussan M, Maisonobe J-A et al. Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. Sci Rep 2014; 55(3): 414–422.
23. van Velden FHP, Kramer GM, Frings V 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; doi:10.1007/s11307-016-0940-2.
24. El Naqa I, Grigsby PW, Apte A et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. Pattern Recognit 2009; 42(6): 1162–1171.
25. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973; SMC3(6): 610–621.
26. Galloway MM. Texture analysis using gray level run lengths. Comput Graph Image Process 1975; 4(2): 172–179.
27. Thibault G, Angulo J, Meyer F. Advanced statistical matrices for texture characterization: application to cell classification. IEEE Trans Biomed Eng. 2014; 61(3): 630–637.

28. Sun C, Wee WG. Neighboring gray level dependence matrix for texture classification. Comput Vis Graph Image Process 1983; 23(3): 341–352.
29. Narang S, Lehrer M, Yang D et al. Radiomics in glioblastoma: current status, challenges and potential opportunities. Transl Cancer Res 2016; 5(4): 383–397.
30. Zheng Y, Englander S, Baloch S et al. STEP: Spatiotemporal enhancement pattern for MR-based breast tumor diagnosis. Med Phys 2009; 36(7): 3192–3204.
31. Wang JJ-Y, Bensmail H, Gao X. Joint learning and weighting of visual vocabulary for bag-of-feature based tissue classification. Pattern Recognit 2013; 46(12): 3249–3255.
32. Cataldo SD, Ficarra E. Mining textural knowledge in biological images: applications, methods and trends. Comput Struct Biotechnol J 2016; doi:10.1016/j.csbj.2016.11.002.
33. Armato SG, Meyer CR, Mcnitt-Gray MF et al. The reference image database to evaluate response to therapy in lung cancer (RIDER) project: a resource for the development of change-analysis software. Clin Pharmacol Ther 2008; 84(4): 448–456.
34. Clarke R, Ressom HW, Wang A et al. The properties of high-dimensional data spaces: implications for exploring gene and protein expression data. Nat Rev Cancer 2008; 8(1): 37–49.
35. Ferte C, Trister AD, Huang E et al. Impact of bioinformatic procedures in the development and translation of high-throughput molecular classifiers in oncology. Clin Cancer Res 2013; 19(16): 4315–4325.
36. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat 1979; 6(2): 65–70.
37. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Statist Med 1990; 9(7): 811–818.
38. Kanas VG, Zacharaki EI, Thomas GA et al. Learning MRI-based classification models for MGMT methylation status prediction in glioblastoma. Comput Methods Programs Biomed 2017; 140: 249–257.
39. Parmar C, Grossmann P, Bussink J et al. Machine learning methods for quantitative radiomic biomarkers. Sci Rep 2015. doi:10.1038/srep13087.
40. Paragios N, Duncan J, Ayache N. Handbook of Biomedical Imaging: Methodologies and Clinical Research, Springer-Verlag New York, Inc. 2015.
41. Kumar D, Shafiee MJ, Chung AG et al. Discovery radiomics for computed tomography cancer detection. 2015; arXiv:1509.00117.
42. Hawkins S, Wang H, Liu Y et al. Predicting malignant nodules from screening CT scans. J Thoracic Oncol 2016; doi:10.1016/j.jtho.2016.07.002.
43. Shen W, Zhou M, Yang F et al. Multi-scale convolutional neural networks for lung nodule classification. In S Ourselin, DC Alexander, C-F Westin, MJ Cardoso (eds): Information Processing in Medical Imaging. Cham: Springer International Publishing, 2015; 9123: 588–599.
44. Zacharaki EI, Morita N, Bhatt P et al. Survival analysis of patients with high-grade gliomas based on data mining of imaging variables. AJNR Am J Neuroradiol. 2012; 33(6): 1065.
45. Zacharaki EI, Kanas VG, Davatzikos C. Investigating machine learning techniques for MRI-based classification of brain neoplasms. Int J Comput Assist Radiol Surg. 2011; 6(6): 821–828.
46. Liu Y, Balagurunathan Y, Atwater T et al. Radiological image traits predictive of cancer status in pulmonary nodules. Clin Cancer Res 2016; clincanres.3102.2016.
47. Zacharaki EI, Wang S, Chawla S et al. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med 2009; 62(6): 1609–1618.
48. Wang J, Liu X, Dong D et al. Prediction of malignant and benign of lung tumor using a quantitative radiomic method. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2016: 1272–1275.
49. Nie K, Shi L, Chen Q et al. Rectal cancer: assessment of neoadjuvant chemo-radiation outcome based on radiomics of multi-parametric MRI. Clin Cancer Res 2016; doi:10.1158/1078-0432.CCR-15-2997.
50. Zhang Q, Xiao Y, Dai W et al. Deep learning based classification of breast tumors with shear-wave elastography. Ultrasonics 2016; 72: 150–157.
51. Setio AAA, Ciompi F, Litjens G et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imaging 2016; 35(5): 1160–1169.
52. Carneiro G, Oakden-Rayner L, Bradley AP et al. Automated 5-year mortality prediction using deep learning and radiomics features from chest computed tomography. 2016.
53. Cheng J-Z, Ni D, Chou Y-H et al. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. Sci Rep 2016; 6: 24454.
54. Kallenberg M, Petersen K, Nielsen M et al. Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring. IEEE Trans Med Imaging 2016; 35(5): 1322–1331.
55. Calhoun PS, Kuszyk BS, Heath DG et al. Three-dimensional volume rendering of spiral CT data: theory and method 1. Radiographics 1999; 19(3): 745–764.
56. Sun R, Orlhac F, Robert C et al. In Regard to Mattonen. Int J Radiat Oncol Biol Phys. 2016; 95(5): 1544–1545.
57. Bentzen SM. Theragnostic imaging for radiation oncology: dose-painting by numbers. Lancet Oncol 2005; 6(2): 112–117.
58. Bentzen SM, Gregoire V. Molecular-imaging-based dose painting – a novel paradigm for radiation therapy prescription. Semin Radiat Oncol 2011; 21(2): 101–110.
59. IMV CT Benchmark Report. IMV Medical Information Division, Inc. 2015.
60. Eskey CJ, Koretsky AP, Domach MM. Jain RK. 2H-nuclear magnetic resonance imaging of tumor blood flow: spatial and temporal heterogeneity in a tissue-isolated mammary adenocarcinoma. Cancer Res 1992; 52(21): 6010–6019.
61. Grove O, Berglund AE, Schabath MB et al. Quantitative computed tomographic descriptors associate tumor shape complexity and intratumor heterogeneity with prognosis in lung adenocarcinoma. PLoS ONE 2015; 10(3): e0118261.
62. Li H, Zhu Y, Burnside ES et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. Npj Breast Cancer 2016; 2: 16012.
63. Ganeshan B, Goh V, Mandeville HC et al. Non-small cell lung cancer: histopathologic correlates for texture parameters at CT. Radiology 2013; 266(1): 326–336.
64. Diehn M, Nardini C, Wang DS et al. Identification of noninvasive imaging surrogates for brain tumor gene-expression modules. Proc Natl Acad Sci 2008; 105(13): 5213–5218.
65. Mazurowski MA, Zhang J, Grimm LJ et al. Radiogenomic analysis of breast cancer: luminal B molecular subtype is associated with enhancement dynamics at MR imaging. Radiology 2014; 273(2): 365–372.
66. Gevaert O, Xu J, Hoang CD et al. Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data—methods and preliminary results. Radiology 2012; 264(2): 387–396.
67. Jamshidi N, Diehn M, Bredel M, Kuo MD. Illuminating radiogenomic characteristics of glioblastoma multiforme through integration of MR imaging, messenger RNA expression, and DNA copy number variation. Radiology 2014; 270(1): 1–2.
68. Hayano K, Tian F, Kambadakone AR et al. Texture analysis of non-contrast-enhanced computed tomography for assessing angiogenesis and survival of soft tissue sarcoma. J Comput Assist Tomogr 2015; 39(4): 607–612.
69. Zhu H, Chen H, Lin Z et al. Identifying molecular genetic features and oncogenic pathways of clear cell renal cell carcinoma through the anatomical (PADUA) scoring system. Oncotarget 2016; 7(9): 10006–10014.
70. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000; 100(1): 57–70.
71. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011; 144(5): 646–674.

72. Ganeshan B, Abaleke S, Young RCD et al. 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(1): 137–143.
73. Panth KM, Leijenaar RTH, Carvalho S 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(3): 462–466.
74. Zhu Y, Li H, Guo W et al. Deciphering genomic underpinnings of quantitative MRI-based radiomic phenotypes of invasive breast carcinoma. Sci Rep 2015; 5: 17787.
75. Zinn PO, Mahajan B, Majadan B et al. Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme. PLoS ONE 2011; 6(10): e25451.
76. Basavanhally AN, Ganesan S, Agner S et al. Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology. IEEE Trans Biomed Eng 2010; 57(3): 642–653.
77. Segal E, Sirlin CB, Ooi C et al. Decoding global gene expression programs in liver cancer by noninvasive imaging. Nat Biotechnol 2007; 25(6): 675–680.
78. Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease, Washington, DC: National Academies Press, 2011.
79. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011; 365(5): 395–409.
80. Zinovev D, Feigenbaum J, Furst J, Raicu D. Probabilistic lung nodule classification with belief decision trees. Conf Proc IEEE Eng Med Biol Soc 2011; 4493–4498.
81. Wu W, Parmar C, Grossmann P et al. Exploratory study to identify radiomics classifiers for lung cancer histology. Front Oncol 2016; doi:10.3389/fonc.2016.00071.
82. Maldonado F, Boland JM, Raghunath S et al. Noninvasive characterization of the histopathologic features of pulmonary nodules of the lung adenocarcinoma spectrum using computer-aided nodule assessment and risk yield (CANARY)—a Pilot study. J Thorac Oncol 2013; 8(4): 452–460.
83. Ganeshan B, Miles KA, Young RCD, 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(1): 101–110.
84. Coroller TP, Grossmann P, Hou Y et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiother Oncol 2015; 114(3): 345–350.
85. Huang Y, Liang C, He L et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol 2016; doi:10.1200/JCO.2015.65.9128.
86. Liang C, Huan Y, He L 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 2014; doi:10.18632/oncotarget.8919.
87. Parmar C, Leijenaar RTH, Grossmann P et al. Radiomic feature clusters and prognostic signatures specific for lung and head & neck cancer. Sci Rep 2015; 5: 11044.
88. Leijenaar RTH, Carvalho S, Hoebers FJP et al. External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma. Acta Oncol 2015; 54(9): 1423–1429.
89. Lee J, Narang S, Martinez J et al. Spatial habitat features derived from multiparametric magnetic resonance imaging data are associated with molecular subtype and 12-month survival status in glioblastoma multiforme. PLoS ONE 2015; 10(9): e0136557.
90. Kickingereder P, Burth S, Wick A et al. Radiomic profiling of glioblastoma: identifying an imaging predictor of patient survival with improved performance over established clinical and radiologic risk models. Radiology.2016; 280(3): 880–889.
91. Liu Y, Kim J, Qu F et al. CT features associated with epidermal growth factor receptor mutation status in patients with lung adenocarcinoma. Radiology 2016; 280(1): 271–280. doi: 10.1148/radiol.2016151455.
92. Aerts HJWL, Grossmann P, Tan Y et al. Defining a radiomic response phenotype: a Pilot study using targeted therapy in NSCLC. Sci Rep 2016; doi:10.1038/srep33860.
93. Michoux N, Van den Broeck S, Lacoste L et al. Texture analysis on MR images helps predicting non-response to NAC in breast cancer. BMC Cancer 2015; 15: 574.
94. 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(1): 89–101.
95. Cook GJR, O'Brien ME, Siddique M et al. Non-small cell lung cancer treated with erlotinib: heterogeneity of (18)F-FDG uptake at PET-association with treatment response and prognosis. Radiology 2015; 276(3): 883–893.
96. Goh V, Ganeshan B, Nathan P et al. Assessment of response to tyrosine kinase inhibitors in metastatic renal cell cancer: CT texture as a predictive biomarker. Radiology.2011; 261(1): 165–171.
97. Bundschuh RA, Dinges J, Neumann L et al. Textural parameters of tumor heterogeneity in ^{18}F-FDG PET/CT for therapy response assessment and prognosis in patients with locally advanced rectal cancer. Sci Rep 2014; 55(6): 891–897.
98. Chiou VL, Burotto M. Pseudoprogression and immune-related response in solid tumors. J Clin Oncol. 2015; 33(31): 3541–3543.
99. Tumeh PC, Harview CL, Yearley JH et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature 2014; 515(7528): 568–571.
100. Ferté C. Dynamic quantitative imaging approaches to identify pseudoprogression in cancer patients treated by immune checkpoints blockers. Sci RepTherapeutics 2015.
101. Infante M, Lutman RF, Imparato S et al. Differential diagnosis and management of focal ground-glass opacities. Eur Respir J 2009; 33(4): 821–827.
102. Mattonen SA, Tetar S, Palma DA et al. Automated texture analysis for prediction of recurrence after stereotactic ablative radiation therapy for lung cancer. Int J Radiat Oncol Biol Phys 2015; 93(3): S5–S6.
103. Antunes J, Viswanath S, Rusu M et al. Radiomics analysis on FLT-PET/MRI for characterization of early treatment response in renal cell carcinoma: a proof-of-concept study. Transl Oncol 2016; 9(2): 155–162.
104. Drucker E, Krapfenbauer K. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. EPMA J 2013; 4(1): 7.
105. Wagner JA, Prince M, Wright EC et al. The Biomarkers Consortium: practice and Pitfalls of open-source precompetitive collaboration. Clin Pharmacol Ther 2010; 87(5): 539–542.
106. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. BMC Med 2012; 10: 51.
107. Brazma A, Hingamp P, Quackenbush J et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet 2001; 29(4): 365–371.
108. O'Connor JPB, Aboagye EO, Adams JE et al. Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol 2016. doi:10.1038/nrclinonc.2016.162.
109. https://www.deepmind.com/health (7 February 2017, date last accessed).
110. Grand-challenges - all challenges. https://grand-challenge.org/All_Challenges/ (7 February 2017, date last accessed).
111. LUNA - Home. https://luna.grand-challenge.org/ (7 February 2017, date last accessed).
112. The Digital Mammography DREAM Challenge – Sage Bionetworks. https://www.synapse.org/#!Synapse:syn4224222/wiki/401743 (7 February 2017, date last accessed).
113. Mackin D, Fave X, Zhang L et al. Measuring computed tomography scanner variability of radiomics features. Invest Radiol 2015; 50(11): 757–765.
114. Kim H, Park CM, Lee M et al. Impact of reconstruction algorithms on CT radiomic features of pulmonary tumors: analysis of intra- and inter-