Review

Clinical Cancer Research

Imaging Intratumor Heterogeneity: Role in Therapy Response, Resistance, and Clinical Outcome

James P.B. O'Connor[1,2], Chris J. Rose[1], John C. Waterton[1,3], Richard A.D. Carano[4], Geoff J.M. Parker[1], and Alan Jackson[1]

Abstract

Tumors exhibit genomic and phenotypic heterogeneity, which has prognostic significance and may influence response to therapy. Imaging can quantify the spatial variation in architecture and function of individual tumors through quantifying basic biophysical parameters such as CT density or MRI signal relaxation rate; through measurements of blood flow, hypoxia, metabolism, cell death, and other phenotypic features; and through mapping the spatial distribution of biochemical pathways and cell signaling networks using PET, MRI, and other emerging molecular imaging techniques. These methods can establish whether one tumor is more or less heterogeneous than another and can identify subregions with differing biology. In this article, we review the image analysis methods currently used to quantify spatial heterogeneity within tumors. We discuss how analysis of intratumor heterogeneity can provide benefit over more simple biomarkers such as tumor size and average function. We consider how imaging methods can be integrated with genomic and pathology data, instead of being developed in isolation. Finally, we identify the challenges that must be overcome before measurements of intratumoral heterogeneity can be used routinely to guide patient care. *Clin Cancer Res; 21(2); 249–57. ©2014 AACR.*

Introduction

Malignant tumors are biologically complex and exhibit substantial spatial variation in gene expression, biochemistry, histopathology, and macroscopic structure. Cancerous cells not only undergo clonal evolution from a single progenitor cell into more aggressive and therapy-resistant cells, but also exhibit branched evolution, whereby each tumor develops and preserves multiple distinct subclonal populations (1). This genetic heterogeneity (1, 2), combined with spatial variation in environmental stressors, leads to regional differences in stromal architecture (3), oxygen consumption (4, 5), glucose metabolism (4), and growth factor expression (6). Consequently, tumor subregions develop, each with spatially distinct patterns of blood flow (7, 8), vessel permeability (9), cell proliferation (10), cell death (11), and other features.

Spatial heterogeneity is found between different tumors within individual patients (intertumor heterogeneity) and within each lesion in an individual (intratumor heterogeneity). Intratumor heterogeneity is near ubiquitous in malignant tumors, but the extent varies between preclinical cancer models and between patients (12). Allowing for these differences, some common themes emerge. First, intratumor heterogeneity can be dynamic. For example, variations in tumor pO_2 fluctuate over minutes to hours (5, 6). Second, intratumor heterogeneity tends to increase as tumors grow (7, 13). Third, established spatial heterogeneity frequently indicates poor clinical prognosis (14), in part due to resistant subpopulations of cells driving resistance to therapy (3, 15). Finally, intratumor heterogeneity may increase or decrease following efficacious anticancer therapy (11, 16), depending on the imaging test used and the underlying tumor biology (17).

Imaging depicts spatial heterogeneity in tumors. However, while imaging is central to diagnosis, staging, response assessment, and recurrence detection in routine oncologic practice, most clinical radiology and research studies only measure tumor size or average parameter values, such as median blood flow (18). In doing so, spatially rich information is discarded. There has been considerable effort to use more sophisticated analyses to either quantify overall tumor spatial complexity or identify the tumor subregions that may drive disease transformation, progression, and drug resistance (11, 19).

In this review, we highlight the strengths and weaknesses of methods that measure intratumor spatial heterogeneity (Fig. 1 and Table 1). We evaluate evidence that heterogeneity analyses provide any clinical benefit over simple "average value" measurements. We discuss how imaging, genomic, and pathology biomarkers of intratumor heterogeneity relate to one another. Finally, we identify the hurdles to translating image biomarkers of spatial heterogeneity into clinical practice.

Qualitative Assessment of Heterogeneity

Radiologists use qualitative descriptors to describe adverse spatial features and functional heterogeneity on clinical scans.

[1]CRUK-EPSRC Cancer Imaging Centre in Cambridge and Manchester, University of Manchester, Manchester, United Kingdom. [2]Department of Radiology, Christie Hospital, Manchester, United Kingdom. [3]R&D Personalised Healthcare and Biomarkers, AstraZeneca, Macclesfield, United Kingdom. [4]Biomedical Imaging Department, Genentech, Inc., South San Francisco, California.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Corresponding Author: James P.B. O'Connor, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, United Kingdom. Phone: 44-1614463896; Fax: 44-1612755145; E-mail: james.o'connor@manchester.ac.uk

doi: 10.1158/1078-0432.CCR-14-0990

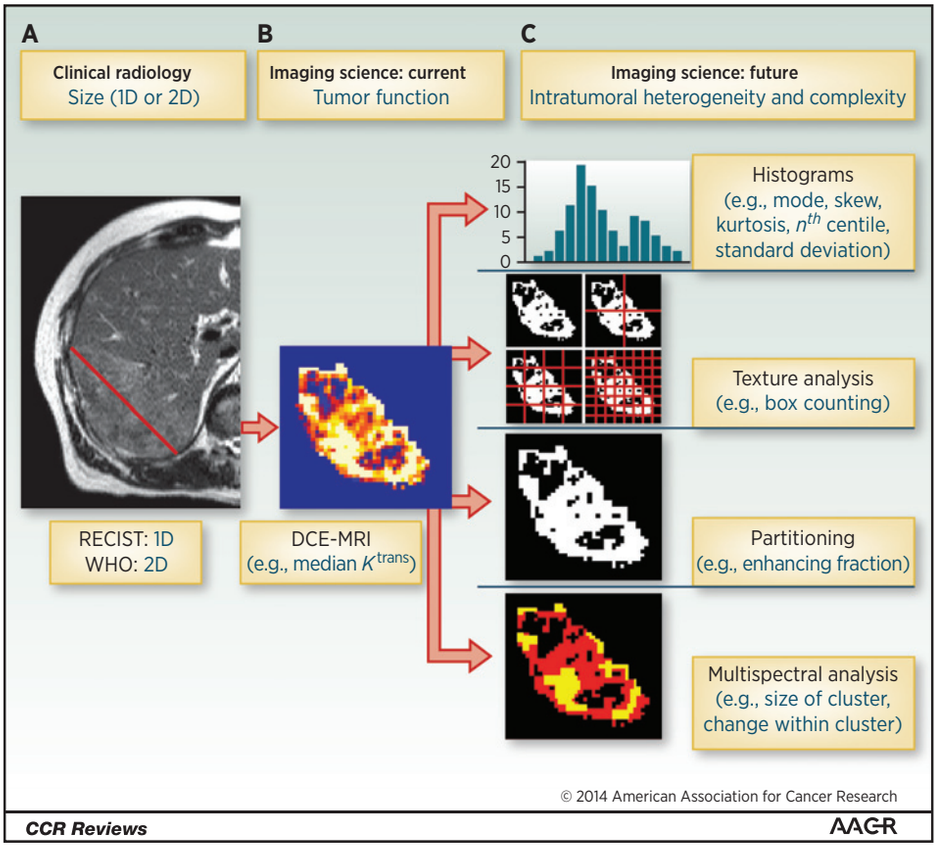

Figure 1.
Quantifying intratumoral heterogeneity: the example liver metastasis from a patient with a colonic primary tumor can be measured in several different ways. A, most clinical assessment of tumors is size based. B, functional imaging methods can measure tumor pathophysiology but tend to derive average parameter values, such as median K^{trans}. C, some intratumoral heterogeneity methods quantify overall complexity of a distribution (histograms) or spatial arrangement of data (texture analysis). Other methods identify tumor subregions using *a priori* assumptions (partitioning) or data-driven approaches (multispectral analysis).

For example, when assessing pulmonary nodules on CT (20) and breast lumps on X-ray mammography (21), spiculation implies greater risk of malignancy compared with well-circumscribed lesions. Indeed, spiculate morphology is part of the BI-RADS (Breast Imaging Reporting and Data System) lexicon that classifies breast lesions as "radiologically malignant" (22).

Identifying a tumor "hot spot" is also commonplace in cancer radiology. The maximum standardized uptake value (SUV_{max}) derived from ^{18}F-FDG PET-CT imaging is an established proxy for identifying abnormal glucose metabolism, based on identifying the one or more voxels with greatest abnormality. Measuring SUV_{max} is simple and reproducible (23, 24) and can be performed on clinical work stations. The presence of abnormal glycolysis in part of a tumor using SUV_{max} is used widely to stage and monitor response in several malignant tumors (25). In glioma, regional high values of tracer uptake (^{18}F-FDG and ^{11}C-methionine) have been used to grade tumors using targeted biopsy (26).

Perfusion CT and MRI methods use "hot spot" analysis to identify tumor regions with the *most* abnormal vascular features. Specialist neuro-oncology centers use dynamic contrast-enhanced MRI (DCE-MRI) or dynamic susceptibility contrast MRI (DSC-MRI) in patients with high-grade glioma (HGG) to map relative cerebral blood volume (rCBV; ref. 27) based on the rationale that "more vascular" regions correspond to highest malignant grade and that this improves prognostic assessment in patients. It is important to note that hot spot analyses are subjective, only identify regions that are maximum (or minimum when gray scale is inverted), and that observer evaluation of heterogeneity is highly influenced by display ranges and color schemes (28) unless objective algorithms are used.

Voxels: Considerations for Quantifying Heterogeneity

Imaging modalities measure biophysical signals in tissues and spatially encode these signals to create clinical images or parameter maps composed of three-dimensional picture elements called voxels. Heterogeneity analyses combine data from many voxels. Several issues arise when interpreting these data.

First, some voxels suffer from partial volume averaging (typically at interface with nontumor tissue). Second, there is inevitable compromise between having sufficient numbers of voxels to perform the analysis versus sufficiently large voxels to overcome noise and keep imaging times practical (29). Most methods of analysis require hundreds to thousands of voxels for robust application. Third, many studies of tumor spatial heterogeneity have used standard clinical data from protocols dictated by clinical rather than research needs. In some cases, sections of tumor were omitted when noncontiguous tumor sampling was used (30), which may confound three-dimensional spatial analyses (31). Fourth, some calculated voxel values, such as apparent diffusion coefficient (ADC), contrast transfer coefficient (K^{trans}), and blood flow, are derived from multiple images obtained over time. The estimation errors associated with motion vary for different parameters and for different voxels, which should be considered when assessing tumor heterogeneity.

Finally, CT, MRI, or PET voxels are usually nonisotropic (slice thickness exceeds in-plane resolution). Dimensions are typically 200 to 2,000 μm for rodent models and 750 to 5,000 μm for clinical tumors. Compared with genomic and histopathology biomarkers, this represents many orders of magnitude difference in scale (32),

Table 1. Examples of beneficial information from analyzing tumor heterogeneity

	Beneficial information from assessment of spatial heterogeneity	
Indication	**Examples of current clinical use (reference)**	**Examples of research application (reference)**
Screening	Subjective morphology (mammography): lesion spiculation used in BI-RADS in breast cancer (22)	Texture analysis: (mammography): improves sensitivity in distinguishing benign and malignant lesions (31)
Diagnosis	Morphology (CT attenuation): lung nodule spiculation (20)	—
Staging (TNM)	Hot spot analysis (SUV_{max}): ^{18}F-FDG PET improves N and M staging in multiple cancers (25)	—
Grading	Hot spot analysis (rCBV): DSC-MRI maps targeted biopsy to accurately grade glioma (27) Hot spot analysis (SUV_{max}): ^{18}F-FDG PET and ^{11}C-methionine maps targeted biopsy to accurately grade glioma (26)	Histogram analysis (rCBV): improves sensitivity and specificity of grading HGG (38)
Early change	—	Histogram analysis (rCBV): detects early transformation of low-grade glioma to HGG (40) Partitioning data (ETV and E_F): assessment of response, reveals subtle changes in tumor biology (48, 72) *A priori* segmentation (K^{trans}): demonstrates differential response in tumor rim and core (62) Data-driven segmentation (K^{trans}): demonstrates differential response in viable and necrotic tumor regions (79)
Outcome	—	Histogram analysis (ADC): data relate to OS (42) Feature analysis (CT and MRI): data relate to PFS and OS in various tumor types (55–59) Partitioning data (ETV and E_F): baseline values prognostic of outcome in HGG (67–69) and cervical cancer (70, 71) Partitioning data (SUV_{max}): persistent values above a threshold indicate poor PFS in GIST or renal cancer treated with TKI (73–75)

Abbreviations: E_F, enhancing voxels; ETV, enhancing tumor volume; GIST, gastrointestinal stromal tumor; OS, overall survival; PFS, progression-free survival; TKI, tyrosine kinase inhibitor; TNM, tumor, nodes, metastasis.

making it difficult to validate image heterogeneity biomarkers against pathology. However, imaging methods have the distinct advantage of whole-tumor sampling, allowing all of the genetic and pathologic variation within tumors to be sampled (33).

Measuring Degree of Heterogeneity: Quantifying Parameter Distributions

Voxel values can be plotted as histograms, from which many simple descriptors can be extracted as potential biomarkers. These include simple descriptors of image heterogeneity such as standard deviation, interquartile range, *n*th centile(s), skew and kurtosis (Supplementary Fig. S1), as well as mean and median values (17, 34). In these approaches, the inherent spatial relationship between voxels is discarded and data are treated as a list of continuous variables.

Histogram analysis

Histograms can be generated using widely available software and have proved popular methods for characterizing intratumoral heterogeneity, accounting for approximately half of all published studies (35). Several important points should be considered. Histogram analyses have high dimensionality and generate many parameters and thus require correction for multiple comparisons (36). The repeatability and reproducibility of many histogram-derived parameters are uncertain and have not yet been evaluated in multicenter studies (29). Furthermore, many parameters, such as fifth centile or kurtosis, have no clear biologic correlate, making biological validation difficult.

Evidence for clinical benefit

More than 200 histogram-based studies have analyzed imaging data of tumor response or outcome (PubMed search performed on October 22, 2014), with rapid rise in numbers recently (Supplementary Fig. S2). Unfortunately, many of these studies were retrospective, performed in small numbers of patients, and did not demonstrate added benefit over simpler measurements of tumor structure and function. For example, in HGG, various histogram parameters have been correlated with overall survival (OS; $P < 0.05$), but relationships were generally equivalent or weaker than those seen with median rCBV values (37). Other studies have also shown marginal benefits. Measurements such as peak height of the CBV histogram had superior sensitivity and negative predictive value over hot spot analysis for distinguishing low-grade glioma from HGG (histogram, 90% sensitivity; hot spot, 55%–76% sensitivity; ref. 38).

The real value in histogram analysis appears two-fold. First, changed heterogeneity in the data distribution may relate to clinical outcome. In cervical cancer, more heterogeneous distributions of FDG-PET voxel values were related to greater risk of lymph node metastases, risk of local recurrence, and worse progression-free survival (PFS; ref. 39). In low-grade glioma, changes in just the top few centiles of voxel enhancement were prognostic for subsequent early transformation into HGG ($P = 0.011$; ref. 40). Preclinical data have shown that regional response to therapy may result in unimodal histograms becoming bimodal (ref. 41; Fig. 2).

Second, histograms can quantify data with complex distributions, where average values may mask important information within the data. For example, in patients with HGG receiving bevacizumab and cytotoxic therapies, pretreatment ADC values had a bimodal distribution. Overall mean ADC did not relate to subsequent PFS ($P = 0.14$), but the mean ADC value of the lower mode related significantly to PFS ($P = 0.004$; ref. 30), suggesting that MRI identified two distinct tumor subregions (Supplementary Fig. S1). These findings have been replicated in a multicenter study

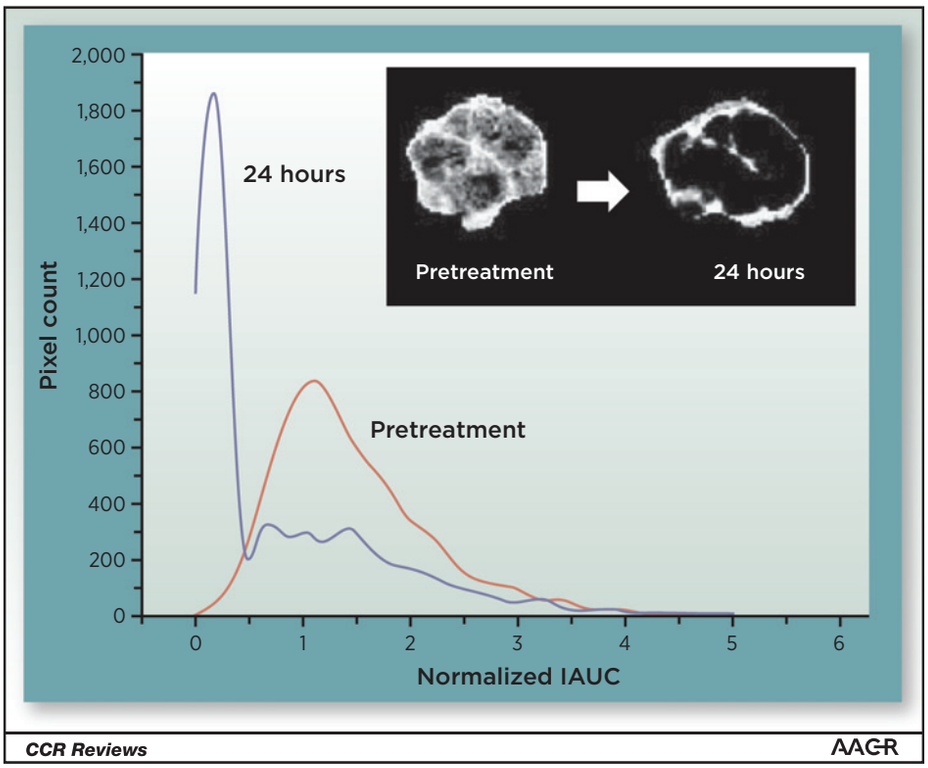

Figure 2.
A GH3 prolactinoma in a rat shows a heterogeneous response to 50 mg/kg of ZD6126. Histogram analysis quantifies this effect and demonstrates transformation from a unimodal distribution to a bimodal one. IAUC, initial area under the contrast agent concentration-time curve. Adapted from Robinson et al. (41).

(42), suggesting possible clinical translation although when histograms from different tumors differ in both location and degree of stretch along the x axis, this can hinder analysis of cohorts of histogram data. Attempts to model for distribution shape may address this problem (43, 44), but remain in their infancy.

Measuring Degree of Heterogeneity: Quantifying Spatial Complexity

Feature analyses measure the spatial complexity of objects. Unlike histogram analysis, the spatial arrangement of voxel values is retained. Related methods, including texture analysis, fractal techniques, and Minkowski functionals, have been applied to tumor data and account for approximately half of studies measuring tumor heterogeneity (35). Initial reports suggest that feature-based metrics derived from ^{18}F-FDG PET (45), ^{18}F-FLT PET (46), CT (47), and MRI (48) have good limits of agreement, with coefficients of variation comparable with summary statistics derived from the same data.

Texture analysis

There is a large body of literature concerning texture analysis in breast cancer. Many studies use the Haralick method (49), where a co-occurrence matrix element denoted $P_{d,\ \theta}\ (i,\ j)$ measures the probability of starting from any image voxel with designated value i, moving d voxels along the image in direction θ, and then arriving at another voxel with value j. This co-occurrence matrix is a two-dimensional histogram describing a joint distribution of all the possible moves with step size d and direction θ on the image. From this, various extracted features, including lesion contrast and correlation, can estimate shape and/or spatial complexity.

Fractal analysis and Minkowski functionals

Fractal dimensions estimate the complexity of geometric patterns resulting from abstract recursive procedures (50). The simplest fractal dimension is the box-counting dimension (d_0), computed by imposing regular grids of a range of scales on a binary object in question and then counting the number of grid elements (boxes) that are occupied by the object at each scale (Supplementary Fig. S3). The box-counting dimension is the slope of the line of best fit when plotting the number of occupied boxes against the reciprocal of the scale on log–log axes (51). Increasingly complex variants can incorporate continuous scale values of parameters such as Hounsfield unit density or K^{trans}. Minkowski functionals analyze binarized images over a range of thresholds and also quantify space-filling properties of tumors.

Evidence for clinical benefit

Texture analyses have been used extensively in X-ray mammography (52). Early applications included discrimination of glandular and fatty regions in mammograms (53) and distinguishing benign and malignant lesions (31). Minkowski functionals analysis can improve stratification of risk for developing breast cancer (receiver operating curve area under curve of >0.9; ref. 54). Here, more heterogeneous images indicated more aggressive tumors.

Feature analyses show the greatest promise as prognostic indicators. CT-based feature analysis parameters predicted OS independent of tumor stage in patients with colorectal cancer treated with cytotoxic therapy ($P < 0.01$; ref. 55), predicted time to progression in patients with renal cancer treated with various antivascular therapies ($P = 0.005$; ref. 56) and predicted OS in non–small cell lung cancer ($P = 0.046$; ref. 57). Similarly, MRI fractal biomarkers have shown prognostic relationships in patients with colorectal cancer treated with bevacizumab

($P < 0.00005$; ref. 58) and in patients with sarcoma treated with cytotoxic therapy (ref. 59; $P < 0.005$). These approaches are now being explored and validated in large populations using an approach termed "radiomics," through which existing clinical imaging data are mined to identify heterogeneity features that predict clinical outcome (60).

Identifying Tumor Subregions

Tumor images contain hundreds to thousands of voxels. Grouping "similar" voxels together (parcellation) may define multiple subregions with common biology that respond differentially to therapy or drive progression. Parcellation techniques differ in their underlying assumptions and methodology. Some use prior knowledge, whereas others rely solely on information contained within the images (Fig. 3).

Parcellation using *a priori* assumptions

Voxels can be categorized by presence or absence of a hallmark, such as enhancement. Alternatively, tumor signals (such as CT density, ^{18}F-FDG SUV_{max}, or K^{trans}) can be parcellated using one or more thresholds. In distinction, tumor regions may be defined geographically, for example by modeling tumors as spheres with concentric radial subregions (61) or by labeling voxels as "rim" or "core" based on relative voxel position in histogram distributions (62).

These simple approaches have important caveats. "Binary" features such as enhancement are not absolute but depend on how images are acquired and analyzed. Threshold values for continuous data based on "cut points" selected to enhance statistical separation in studies are arbitrary. *A priori* methods may have a misleading biologic basis. The Macdonald criteria in HGG illustrate this point, as measuring contrast-enhancing tumor only has been superseded by the Response Assessment in Neuro-Oncology criteria, which incorporate measures of nonenhancing tissue (with likely microscopic foci of infiltrative neoplastic cells) into response criteria (63).

Data-driven segmentation

Some studies acquire multiple imaging parameters (for example, K^{trans}, ADC, and ^{18}F-FDG SUV_{max}) and analyze each signal independently (64). An alternative strategy parcellates voxels with similar signals (or "spectra") into functionally coherent regions within a lesion (65). Most multispectral analyses use pattern recognition techniques that simultaneously analyze images to identify voxel clusters in a multidimensional feature space. A classifier then groups individual voxels together based on their similarities and differences (66).

Evidence for clinical benefit

Multiple studies of patients with HGG have shown lower baseline lesion enhancing tumor volume (ETV) had beneficial OS ($P = 0.0026$; ref. 67) or PFS ($P = 0.0309$; ref. 68) and that early reduction in ETV after bevacizumab related to OS ($P = 0.0008$) (69), whereas WTV and K^{trans} did not. Studies of solid tumors outside the brain have shown that ETV or the proportion of enhancing voxels (E_F) before treatment is correlated with PFS in cervical cancer (70, 71). Trials of antiangiogenic drugs suggest that E_F provides pharmacodynamic information, independent from other DCE-MRI parameters (48, 72). These data support the hypothesis that measuring tumor regions may be a more useful biomarker than average values of whole tumors in some clinical scenarios.

Threshold-derived partitioning has shown value in patients with gastrointestinal stromal tumors (GIST) imaged with

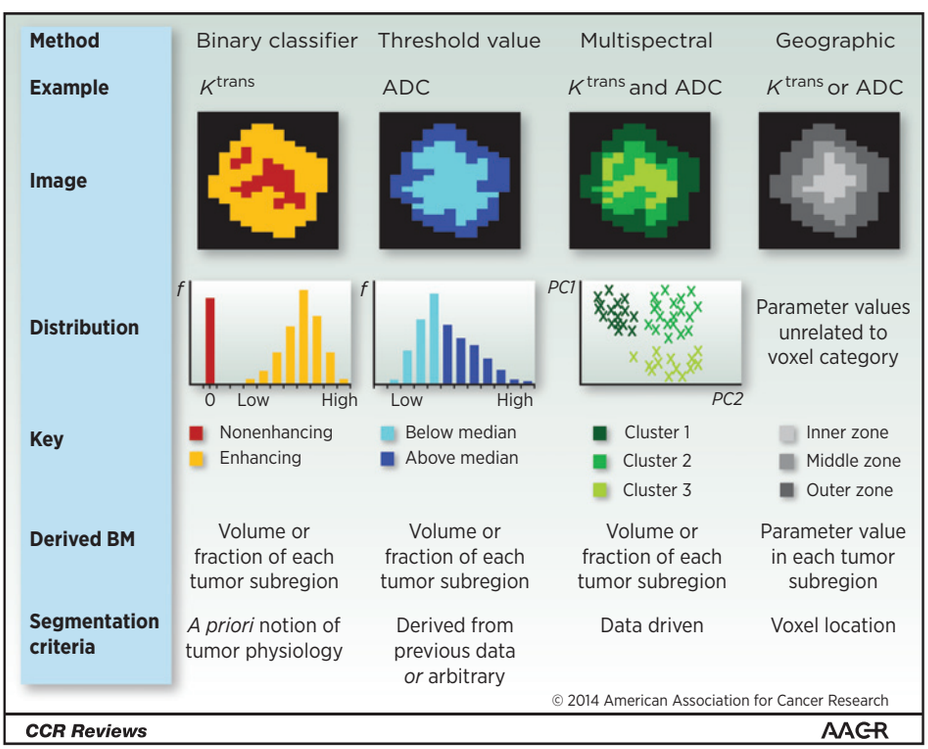

Figure 3.
Quantifying tumor subregions: *a priori* methods parcellate tumor regions based on binary classifiers [e.g., nonenhancing (= 0) or enhancing (= 1) voxels from a K^{trans} map] or threshold classification using a "cut point" defined from previous data, or an arbitrary value such as the median (e.g., of an ADC distribution). Multispectral classification generates clusters based on data-driven segmentation from multiple signals (e.g., K^{trans} and ADC), for example using principal components (PC) analysis (graph shows three clusters based on PC1 and PC2). The volume or proportional fraction of each region is measured in these methods. Geographic methods are distinct because they define subregions based on the location of a pixel rather than its parameter value(s). BM, biomarker.

^{18}F-FDG PET-CT. Here, tumor SUV below 8 g/mL 4 weeks after treatment with sunitinib was associated with markedly longer PFS than those with SUV above the same level (29 vs. 4 weeks; $P < 0.0001$; ref. 73). Similarly, SUV thresholds have distinguished good and poor time to treatment failure in GIST patients treated with imatinib ($P < 0.0001$; ref. 74). Furthermore, there is some evidence that the presence of high SUV_{max} values may relate to OS in some cancers (75). In these studies, cutoff points have been chosen using *post hoc* criteria. Prospective studies are required to validate these thresholds in specific patient–therapy combinations if these techniques are to have clinical translation.

Data-driven approaches have successfully distinguished viable tumor from nonviable tumor using multiparametric MRI and validated the method against H&E histology (76, 77). Moreover, image-defined regions of viable and nonviable tumor show differential response to radiotherapy (78) and to anti-VEGF antibody (ref. 79; Fig. 4). Multispectral analyses of baseline data that identify tumor subregions with distinct biology, responsible for driving tumor response, resistance, and progression, are highly attractive, although performed rarely. Algorithms to use for this type of analysis are commercially available (80).

Unfortunately, it is very difficult to track *change* in individual voxels. Some progress has been made in studies of HGG using a method termed "response parametric mapping," in which rCBV and ADC maps have been categorized as showing no change or increasing/decreasing by greater than 20% following therapy. The proportion of tumor with reduced rCBV was correlated with OS ($P = 0.019$), where mean rCBV was not (81). However, it is extremely difficult to extend this approach to accommodate changes in tumor size, orientation, and deformation, making it difficult to translate these methods from specialist laboratories to health care systems.

Integrating Imaging, Genomics, and Histopathology

Recent work highlights the intratumor variation in gene mutation and expression (33), but few studies have explored the spatial relationship among imaging, genomics, and histopathology. Preclinical studies have reported differential levels of gene activation and protein transcription that relate to regional perfusion as measured with DCE-MRI (82) and related differential gene expression to regional PET signals (83). Pilot data have associated heterogeneous enhancement patterns with genetic subtypes of breast cancer (84). The ADC characteristics at tumor margins, rather than overall mean values, have been shown to correctly categorize oligodendrogliomas by their 1p/19q loss status (85).

These data highlight the need for larger prospective studies to elucidate how and when integration of imaging, genomic, and pathology data may be useful. These studies must evaluate large complex data across a range of biologically different scales (15, 32). Two questions require urgent answers. First, do imaging, genomics, and histopathology show spatial correspondence because they *measure the same biology* in different ways, and if so,

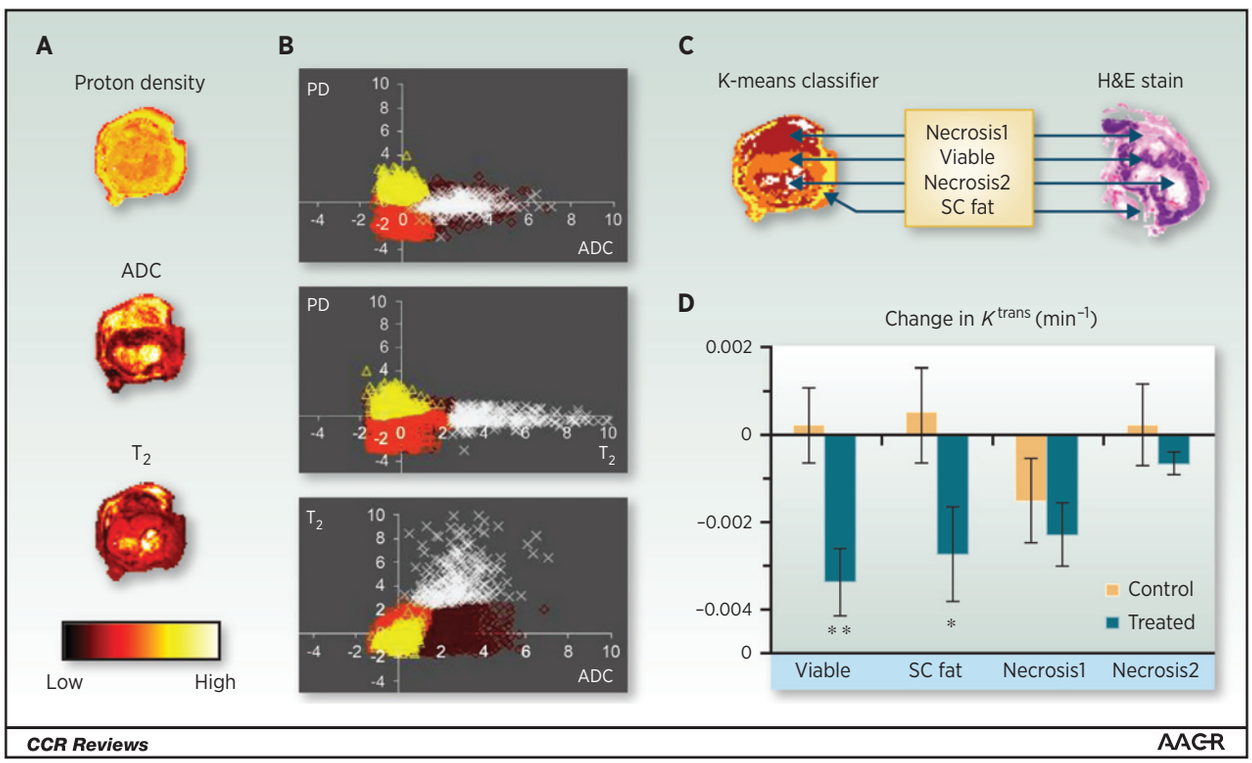

Figure 4.
Multispectral analysis. A, ADC, T_2 values and proton density (PD) maps shown for a colorectal cancer murine xenograft. B, K-means clustering is performed and decision boundaries are formed. C, four tissue categories are produced [viable tumor, subcutaneous (SC) fat, and two necrotic regions (differentiated from each other by the T_2 signal)] that relate to pathology as seen on the H&E section. D, the antivascular effect of G6-31 in treated animals is significant in viable tumor tissue but not significant in necrotic regions. Adapted from Carano et al. (76) and Berry et al. (79).

Table 2. Image analysis methods for quantifying spatial heterogeneity: comparison of strengths and weaknesses

Analysis	Availability	Literature	Spatial information	Single-center reproducibility	Potential statistical and design issues	Biologic correlate
Qualitative interpretation by radiologist						
Subjective morphology	Worldwide	Extensive	Evaluated qualitatively	Largely unknown	Subjective and difficult to apply in quantitative study	Measure of overall complexity
Hot spot	Worldwide	Extensive	Ignored outside hot spot	Very good[a]	Subjective	Difficult to evaluate
Quantitative measurement by scientist						
Histogram analysis	Extensive	Extensive	Discarded	Largely unknown	Large number of parameters derived: multiple comparisons	Difficult to evaluate
Feature analysis	Specialist technique	Moderate	Used to compute spatial complexity	Very good[a]	Large number of parameters derived: multiple comparisons	Measure of overall complexity
A priori segmentation	Specialist technique	Small	Used to define subregions	Largely unknown	Various methods; category definition may be arbitrary	Subregion identified: can match to histology
Multispectral analysis	Specialist technique	Minimal	Used to define subregions	Unknown	Multidimensional datasets	Subregion identified: can match to histology

[a]In some studies.

can imaging reliably identify important subregions noninvasively? If so, imaging could extend personalized medicine to tumors that cannot be biopsied safely. Alternatively, do imaging, genomics, and histopathology *measure different biology*, providing complementary data? If so, this would open up new multidisciplinary strategies for advancing personalized medicine focusing phenotype and genome together, rather than genome in isolation.

Future Directions and Conclusions

Diverse philosophical and mathematical approaches can quantify intratumor heterogeneity. Most techniques can be applied across all imaging modalities, while some require datasets with multiple signals. Each method has strengths and limitations (Table 2). Advances in hardware such as simultaneous PET–MRI and whole-body imaging with MRI diffusion will further stimulate imaging research into analysis of both intratumor heterogeneity and in differences between multiple lesions within individuals.

All methods must overcome significant hurdles before cementing a role in cancer radiology. First, clear patient benefit must be demonstrated from heterogeneity analyses, above and beyond that found with simple imaging (and nonimaging) biomarkers. Second, promising data must be replicated in other institutions and biomarkers must be validated (35, 86). Third, relationships between heterogeneity parameters and underlying biology must be established. Finally, the development of large imaging datasets with multiple candidate "probes" has become an area of intense interest, attempting to parallel advances in next-generation sequencing. This "radiomic" approach (60) places heterogeneity parameters at its center, but the sheer number of parameters under investigation incurs statistical dangers due to large numbers of multiple comparisons. As these challenges are addressed, heterogeneity analyses have great potential to translate from the research arena to play a role in clinical decision making.

Disclosure of Potential Conflicts of Interest

J.C. Waterton is an employee of and has ownership interest (including patents) in AstraZeneca. R.A.D. Carano is an employee of Genentech and has ownership interest (including patents) in Roche. G.J.M. Parker is CEO of and has ownership interest (including patents) in Bioxydyn Limited; reports receiving commercial research grants from AstraZeneca, Merck Serono, and Roche; and is a consultant/advisory board member for GlaxoSmithKline. No potential conflicts of interest were disclosed by the other authors.

Grant Support

J.P.B. O'Connor is supported by a Cancer Research UK Clinician Scientist Fellowship (C19221/A15267). J.P.B. O'Connor, G.J.M. Parker, and A. Jackson are supported by a Cancer Research UK and EPSRC Comprehensive Imaging Centre grant (C8742/A018097).

Received May 27, 2014; revised October 22, 2014; accepted October 28, 2014; published OnlineFirst November 24, 2014.

References

1. Swanton C. Intratumor heterogeneity: evolution through space and time. Cancer Res 2012;72:4875–82.
2. Heppner GH. Tumor heterogeneity. Cancer Res 1984;44:2259–65.
3. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. Nature 2013;501:346–54.
4. Schroeder T, Yuan H, Viglianti BL, Peltz C, Asopa S, Vujaskovic Z, et al. Spatial heterogeneity and oxygen dependence of glucose consumption in R3230Ac and fibrosarcomas of the Fischer 344 rat. Cancer Res 2005;65:5163–71.
5. Cardenas-Navia LI, Mace D, Richardson RA, Wilson DF, Shan S, Dewhirst MW. The pervasive presence of fluctuating oxygenation in tumors. Cancer Res 2008;68:5812–9.
6. Serganova I, Doubrovin M, Vider J, Ponomarev V, Soghomonyan S, Beresten T, et al. Molecular imaging of temporal dynamics and spatial heterogeneity of hypoxia-inducible factor-1 signal transduction activity in tumors in living mice. Cancer Res 2004;64:6101–8.
7. 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;52:6010–9.
8. Hamberg LM, Kristjansen PE, Hunter GJ, Wolf GL, Jain RK. Spatial heterogeneity in tumor perfusion measured with functional computed tomography at 0.05 microliter resolution. Cancer Res 1994;54:6032–6.
9. Degani H, Gusis V, Weinstein D, Fields S, Strano S. Mapping pathophysiological features of breast tumors by MRI at high spatial resolution. Nat Med 1997;3:780–2.
10. Choi YP, Shim HS, Gao MQ, Kang S, Cho NH. Molecular portraits of intratumoral heterogeneity in human ovarian cancer. Cancer Lett 2011;307:62–71.
11. Gatenby RA, Silva AS, Gillies RJ, Frieden BR. Adaptive therapy. Cancer Res 2009;69:4894–903.
12. Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. Nature 2013;501:328–37.

13. Brurberg KG, Gaustad JV, Mollatt CS, Rofstad EK. Temporal heterogeneity in blood supply in human tumor xenografts. Neoplasia 2008;10: 727–35.
14. Shipitsin M, Campbell LL, Argani P, Weremowicz S, Bloushtain-Qimron N, Yao J, et al. Molecular definition of breast tumor heterogeneity. Cancer Cell 2007;11:259–73.
15. Gatenby RA, Grove O, Gillies RJ. Quantitative imaging in cancer evolution and ecology. Radiology 2013;269:8–15.
16. Simpson-Herren L, Noker PE, Wagoner SD. Variability of tumor response to chemotherapy. II. Contribution of tumor heterogeneity. Cancer Chemother Pharmacol 1988;22:131–6.
17. Just N. Improving tumour heterogeneity MRI assessment with histograms. Br J Cancer 2014;111:2205–13.
18. O'Connor JP, Jackson A, Asselin MC, Buckley DL, Parker GJ, Jayson GC. Quantitative imaging biomarkers in the clinical development of targeted therapeutics: current and future perspectives. Lancet Oncol 2008;9:766–76.
19. Gonzalez-Garcia I, Sole RV, Costa J. Metapopulation dynamics and spatial heterogeneity in cancer. Proc Natl Acad Sci U S A 2002;99:13085–9.
20. Xu DM, van Klaveren RJ, de Bock GH, Leusveld A, Zhao Y, Wang Y, et al. Limited value of shape, margin and CT density in the discrimination between benign and malignant screen detected solid pulmonary nodules of the NELSON trial. Eur J Radiol 2008;68:347–52.
21. Burrell HC, Pinder SE, Wilson AR, Evans AJ, Yeoman LJ, Elston CW, et al. The positive predictive value of mammographic signs: a review of 425 nonpalpable breast lesions. Clin Radiol 1996;51:277–81.
22. Orel SG, Kay N, Reynolds C, Sullivan DC. BI-RADS categorization as a predictor of malignancy. Radiology 1999;211:845–50.
23. Huang YE, Chen CF, Huang YJ, Konda SD, Appelbaum DE, Pu Y. Interobserver variability among measurements of the maximum and mean standardized uptake values on (18)F-FDG PET/CT and measurements of tumor size on diagnostic CT in patients with pulmonary tumors. Acta Radiol 2010;51:782–8.
24. Jackson T, Chung MK, Mercier G, Ozonoff A, Subramaniam RM. FDG PET/CT interobserver agreement in head and neck cancer: FDG and CT measurements of the primary tumor site. Nucl Med Commun 2012;33:305–12.
25. Fletcher JW, Djulbegovic B, Soares HP, Siegel BA, Lowe VJ, Lyman GH, et al. Recommendations on the use of 18F-FDG PET in oncology. J Nucl Med 2008;49:480–508.
26. Pirotte B, Goldman S, Massager N, David P, Wikler D, Vandesteene A, et al. Comparison of 18F-FDG and 11C-methionine for PET-guided stereotactic brain biopsy of gliomas. J Nucl Med 2004;45:1293–8.
27. Knopp EA, Cha S, Johnson G, Mazumdar A, Golfinos JG, Zagzag D, et al. Glial neoplasms: dynamic contrast-enhanced T2*-weighted MR imaging. Radiology 1999;211:791–8.
28. O'Connor JP, Jackson A, Parker GJ, Roberts C, Jayson GC. Dynamic contrast-enhanced MRI in clinical trials of antivascular therapies. Nat Rev Clin Oncol 2012;9:167–77.
29. Jackson A, O'Connor JP, Parker GJ, Jayson GC. Imaging tumor vascular heterogeneity and angiogenesis using dynamic contrast-enhanced magnetic resonance imaging. Clin Cancer Res 2007;13:3449–59.
30. Pope WB, Kim HJ, Huo J, Alger J, Brown MS, Gjertson D, et al. Recurrent glioblastoma multiforme: ADC histogram analysis predicts response to bevacizumab treatment. Radiology 2009;252:182–9.
31. 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.
32. Cebulla J, Kim E, Rhie K, Zhang J, Pathak AP. Multiscale and multimodality visualization of angiogenesis in a human breast cancer model. Angiogenesis 2014;17:695–709.
33. 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.
34. Issa B, Buckley DL, Turnbull LW. Heterogeneity analysis of Gd-DTPA uptake: improvement in breast lesion differentiation. J Comput Assist Tomogr 1999;23:615–21.
35. Alic L, Niessen WJ, Veenland JF. Quantification of heterogeneity as a biomarker in tumor imaging: a systematic review. PLoS One 2014;9: e110300.
36. Yang X, Knopp MV. Quantifying tumor vascular heterogeneity with dynamic contrast-enhanced magnetic resonance imaging: a review. J Biomed Biotechnol 2011;2011:732848.
37. Law M, Young R, Babb J, Pollack E, Johnson G. Histogram analysis versus region of interest analysis of dynamic susceptibility contrast perfusion MR imaging data in the grading of cerebral gliomas. AJNR Am J Neuroradiol 2007;28:761–6.
38. Emblem KE, Nedregaard B, Nome T, Due-Tonnessen P, Hald JK, Scheie D, et al. Glioma grading by using histogram analysis of blood volume heterogeneity from MR-derived cerebral blood volume maps. Radiology 2008;247:808–17.
39. Kidd EA, Grigsby PW. Intratumoral metabolic heterogeneity of cervical cancer. Clin Cancer Res 2008;14:5236–41.
40. Tofts PS, Benton CE, Weil RS, Tozer DJ, Altmann DR, Jager HR, et al. Quantitative analysis of whole-tumor Gd enhancement histograms predicts malignant transformation in low-grade gliomas. J Magn Reson Imaging 2007;25:208–14.
41. Robinson SP, McIntyre DJ, Checkley D, Tessier JJ, Howe FA, Griffiths JR, et al. Tumour dose response to the antivascular agent ZD6126 assessed by magnetic resonance imaging. Br J Cancer 2003;88:1592–7.
42. Pope WB, Qiao XJ, Kim HJ, Lai A, Nghiemphu P, Xue X, et al. Apparent diffusion coefficient histogram analysis stratifies progression-free and overall survival in patients with recurrent GBM treated with bevacizumab: a multi-center study. J Neurooncol 2012;108:491–8.
43. O'Connor E, Fieller N, Holmes A, Waterton JC, Ainscow E. Functional principal component analyses of biomedical images as outcome measures. J R Stat Soc (Ser C Appl Stat) 2010;59:57–76.
44. Rose CJ, O'Connor JPB, Cootes TF, Taylor CJ, Jayson GC, Parker GJM, et al. Indexed distribution analysis for improved significance testing of spatially heterogeneous parameter maps: application to dynamic contrast-enhanced MRI biomarkers. Magn Reson Med 2014;71:1299–311.
45. Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. J Nucl Med 2012;53:693–700.
46. Willaime JM, Turkheimer FE, Kenny LM, Aboagye EO. Quantification of intra-tumour cell proliferation heterogeneity using imaging descriptors of 18F fluorothymidine-positron emission tomography. Phys Med Biol 2013;58:187–203.
47. Sanghera B, Banerjee D, Khan A, Simcock I, Stirling JJ, Glynne-Jones R, et al. Reproducibility of 2D and 3D fractal analysis techniques for the assessment of spatial heterogeneity of regional blood flow in rectal cancer. Radiology 2012;263:865–73.
48. O'Connor JP, Carano RA, Clamp AR, Ross J, Ho CC, Jackson A, et al. Quantifying antivascular effects of monoclonal antibodies to vascular endothelial growth factor: insights from imaging. Clin Cancer Res 2009; 15:6674–82.
49. Haralick RM, Shanmuga K, Dinstein I. Textural features for image classification. IEEE Trans Systems Man Cybernetics 1973;6:610–21.
50. Peitgen H-O, Jurgens H, Saupe D. Chaos and fractals. Berlin: Springer; 2004.
51. Rose CJ, Mills SJ, O'Connor JP, Buonaccorsi GA, Roberts C, Watson Y, et al. Quantifying spatial heterogeneity in dynamic contrast-enhanced MRI parameter maps. Magn Reson Med 2009;62:488–99.
52. Oliver A, Freixenet J, Marti J, Perez E, Pont J, Denton ER, et al. A review of automatic mass detection and segmentation in mammographic images. Med Image Anal 2010;14:87–110.
53. Miller P, Astley S. Classification of breast tissue by texture analysis. Image Vis Comput 1992;10:277–82.
54. Li H, Giger ML, Olopade OI, Lan L. Fractal analysis of mammographic parenchymal patterns in breast cancer risk assessment. Acad Radiol 2007; 14:513–21.
55. Ng F, Ganeshan B, Kozarski R, Miles KA, Goh V. Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival. Radiology 2013;266:177–84.
56. Goh V, Ganeshan B, Nathan P, Juttla JK, Vinayan A, Miles KA. Assessment of response to tyrosine kinase inhibitors in metastatic renal cell cancer: CT texture as a predictive biomarker. Radiology 2011;261:165–71.
57. Balagurunathan Y, Gu Y, Wang H, Kumar V, Grove O, Hawkins S, et al. Reproducibility and prognosis of quantitative features extracted from CT images. Transl Oncol 2014;7:72–87.
58. 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.
59. Alic L, van Vliet M, van Dijke CF, Eggermont AM, Veenland JF, Niessen WJ. Heterogeneity in DCE-MRI parametric maps: a biomarker for treatment response? Phys Med Biol 2011;56:1601–16.
60. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
61. Gaustad JV, Benjaminsen IC, Graff BA, Brurberg KG, Ruud EB, Rofstad EK. Intratumor heterogeneity in blood perfusion in orthotopic human melanoma xenografts assessed by dynamic contrast-enhanced magnetic resonance imaging. J Magn Reson Imaging 2005;21:792–800.
62. Checkley D, Tessier JJ, Kendrew J, Waterton JC, Wedge SR. Use of dynamic contrast-enhanced MRI to evaluate acute treatment with ZD6474, a VEGF signalling inhibitor, in PC-3 prostate tumours. Br J Cancer 2003;89:1889–95.
63. Wen PY, Macdonald DR, Reardon DA, Cloughesy TF, Sorensen AG, Galanis E, et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. J Clin Oncol 2010;28:1963–72.
64. Dingemans AM, de Langen AJ, van den Boogaart V, Marcus JT, Backes WH, Scholtens HT, et al. First-line erlotinib and bevacizumab in patients with locally advanced and/or metastatic non-small-cell lung cancer: a phase II study including molecular imaging. Ann Oncol 2011;22:559–66.
65. Padhani AR, Miles KA. Multiparametric imaging of tumor response to therapy. Radiology 2010;256:348–64.
66. Vannier MW, Butterfield RL, Jordan D, Murphy WA, Levitt RG, Gado M. Multispectral analysis of magnetic resonance images. Radiology 1985; 154:221–4.
67. Zhang W, Kreisl TN, Solomon J, Reynolds RC, Glen DR, Cox RW, et al. Acute effects of bevacizumab on glioblastoma vascularity assessed with DCE-MRI and relation to patient survival. Proceedings ISMRM 2009;17:282.
68. Ellingson BM, Cloughesy TF, Lai A, Nghiemphu PL, Mischel PS, Pope WB. Quantitative volumetric analysis of conventional MRI response in recurrent glioblastoma treated with bevacizumab. Neuro Oncol 2011;13:401–9.
69. Kreisl TN, Zhang W, Odia Y, Shih JH, Butman JA, Hammoud D, et al. A phase II trial of single-agent bevacizumab in patients with recurrent anaplastic glioma. Neuro Oncol 2011;13:1143–50.
70. Donaldson SB, Buckley DL, O'Connor JP, Davidson SE, Carrington BM, Jones AP, et al. Enhancing fraction measured using dynamic contrast-enhanced MRI predicts disease-free survival in patients with carcinoma of the cervix. Br J Cancer 2010;102:23–6.
71. Mannelli L, Patterson AJ, Zahra M, Priest AN, Graves MJ, Lomas DJ, et al. Evaluation of nonenhancing tumor fraction assessed by dynamic contrast-enhanced MRI subtraction as a predictor of decrease in tumor volume in response to chemoradiotherapy in advanced cervical cancer. AJR Am J Roentgenol 2010;195:524–7.
72. Messiou C, Orton M, Ang JE, Collins DJ, Morgan VA, Mears D, et al. Advanced solid tumors treated with cediranib: comparison of dynamic contrast-enhanced MR imaging and CT as markers of vascular activity. Radiology 2012;265:426–36.
73. Prior JO, Montemurro M, Orcurto MV, Michielin O, Luthi F, Benhattar J, et al. Early prediction of response to sunitinib after imatinib failure by 18F-fluorodeoxyglucose positron emission tomography in patients with gastrointestinal stromal tumor. J Clin Oncol 2009;27:439–45.
74. Holdsworth CH, Badawi RD, Manola JB, Kijewski MF, Israel DA, Demetri GD, et al. CT and PET: early prognostic indicators of response to imatinib mesylate in patients with gastrointestinal stromal tumor. AJR Am J Roentgenol 2007;189:W324–30.
75. Kayani I, Avril N, Bomanji J, Chowdhury S, Rockall A, Sahdev A, et al. Sequential FDG-PET/CT as a biomarker of response to Sunitinib in metastatic clear cell renal cancer. Clin Cancer Res 2011;17:6021–8.
76. Carano RA, Ross AL, Ross J, Williams SP, Koeppen H, Schwall RH, et al. Quantification of tumor tissue populations by multispectral analysis. Magn Reson Med 2004;51:542–51.
77. Henning EC, Azuma C, Sotak CH, Helmer KG. Multispectral quantification of tissue types in a RIF-1 tumor model with histological validation. Part I. Magn Reson Med 2007;57:501–12.
78. Henning EC, Azuma C, Sotak CH, Helmer KG. Multispectral tissue characterization in a RIF-1 tumor model: monitoring the ADC and T2 responses to single-dose radiotherapy. Part II. Magn Reson Med 2007;57:513–9.
79. Berry LR, Barck KH, Go MA, Ross J, Wu X, Williams SP, et al. Quantification of viable tumor microvascular characteristics by multispectral analysis. Magn Reson Med 2008;60:64–72.
80. Chen L, Choyke PL, Chan TH, Chi CY, Wang G, Wang Y. Tissue-specific compartmental analysis for dynamic contrast-enhanced MR imaging of complex tumors. IEEE Trans Med Imaging 2011;30:2044–58.
81. Galban CJ, Chenevert TL, Meyer CR, Tsien C, Lawrence TS, Hamstra DA, et al. Prospective analysis of parametric response map-derived MRI biomarkers: identification of early and distinct glioma response patterns not predicted by standard radiographic assessment. Clin Cancer Res 2011;17: 4751–60.
82. Costouros NG, Lorang D, Zhang Y, Miller MS, Diehn FE, Hewitt SM, et al. Microarray gene expression analysis of murine tumor heterogeneity defined by dynamic contrast-enhanced MRI. Mol Imaging 2002;1:301–8.
83. von Forstner C, Egberts JH, Ammerpohl O, Niedzielska D, Buchert R, Mikecz P, et al. Gene expression patterns and tumor uptake of 18F-FDG, 18F-FLT, and 18F-FEC in PET/MRI of an orthotopic mouse xenotransplantation model of pancreatic cancer. J Nucl Med 2008;49:1362–70.
84. Yamamoto S, Maki DD, Korn RL, Kuo MD. Radiogenomic analysis of breast cancer using MRI: a preliminary study to define the landscape. AJR Am J Roentgenol 2012;199:654–63.
85. Jenkinson MD, Smith TS, Brodbelt AR, Joyce KA, Warnke PC, Walker C. Apparent diffusion coefficients in oligodendroglial tumors characterized by genotype. J Magn Reson Imaging 2007;26:1405–12.
86. Waterton JC, Pylkkanen L. Qualification of imaging biomarkers for oncology drug development. Eur J Cancer 2012;48:409–15.