



Review

Radiogenomic Analysis of Oncological Data: A Technical Survey

Mariarosaria Incoronato ^{*,†}, Marco Aiello [†], Teresa Infante, Carlo Cavaliere,
Anna Maria Grimaldi, Peppino Mirabelli, Serena Monti and Marco Salvatore

IRCCS SDN, Via E. Gianturco, 113, 80143 Naples, Italy; maiello@sdn-napoli.it (M.A.);
tinfante@sdn-napoli.it (T.I.); ccavaliere@sdn-napoli.it (C.C.); agrimaldi@sdn-napoli.it (A.M.G.);
pmirabelli@sdn-napoli.it (P.M.); smonti@sdn-napoli.it (S.M.); direzionescientifica@sdn-napoli.it (M.S.)

* Correspondence: mincoronato@sdn-napoli.it; Tel.: +39-081-2408260; Fax: +39-081-668841

† These authors contributed equally to this work.

Academic Editors: Jamal Zweit and Sundaresan Gobalakrishnan

Received: 30 December 2016; Accepted: 8 April 2017; Published: 12 April 2017

Abstract: In the last few years, biomedical research has been boosted by the technological development of analytical instrumentation generating a large volume of data. Such information has increased in complexity from basic (i.e., blood samples) to extensive sets encompassing many aspects of a subject phenotype, and now rapidly extending into genetic and, more recently, radiomic information. Radiogenomics integrates both aspects, investigating the relationship between imaging features and gene expression. From a methodological point of view, radiogenomics takes advantage of non-conventional data analysis techniques that reveal meaningful information for decision-support in cancer diagnosis and treatment. This survey is aimed to review the state-of-the-art techniques employed in radiomics and genomics with special focus on analysis methods based on molecular and multimodal probes. The impact of single and combined techniques will be discussed in light of their suitability in correlation and predictive studies of specific oncologic diseases.

Keywords: radiogenomics; cancer; MR; texture analysis; microarray; NGS technologies; correlation matrix; molecular imaging; data mining

1. Introduction

Cancer diagnosis and classification are traditionally based on the histological examination of bioptic specimens. However, this approach presents challenges related to the invasive gathering of tissue, failures to distinguish between clinically relevant subtypes of cancer, and inter- and intra-observer variability [1]. In response to these drawbacks, new high-throughput platforms have emerged with the aim of better characterizing cancer at the molecular level, allowing an earlier diagnosis, better stratification, and more accurate prognosis than the histopathological approaches for targeted treatments. In this scenario, technological improvements in the field of imaging and molecular biology have led to “radiogenomics” or “imaging genomics” [2]. Literally, radiogenomics refers to the analytical processes aimed at correlating cancer imaging features (radiomics) with genomic data (genomics) [3].

The concept behind radiogenomics is the possibility of investigating the relationship between imaging, genomics, and clinical knowledge simply by looking at data, regardless of any qualitative interpretation; roughly speaking, by letting the data speak for themselves [4,5]. Therefore, radiogenomic approaches are extensively based on numerical calculus and computer science methods, allowing the management and analysis of a very large number of variables for each sample and modality. For instance, the image of an oncological lesion (as shown in Figure 1) is commonly rated by a radiologist using functional/morphological attributes (spiculated, enhanced, hyper-intense, focal,

glucose avid), but the imaging study, itself, is a multivariate source of data; therefore, the main aim of radiomics is to extract meaningful information directly from data. How much radiogenomics has benefited from data processing techniques is clear, as it requires sophisticated algorithms and high-performance computing for the analytical processes. This latter consideration reveals the inherent multi-disciplinary nature of this emerging field, where radiologists, oncologists, biologists, and computer and statistical scientists work together to gain useful insights for personalized medicine. The wide development of this novel analytical approach is not limited to oncology; it is also successfully employed in neurology, where textural features related to oncological lesions can be replaced by morpho-functional features or connectivity of brain structures [6,7].

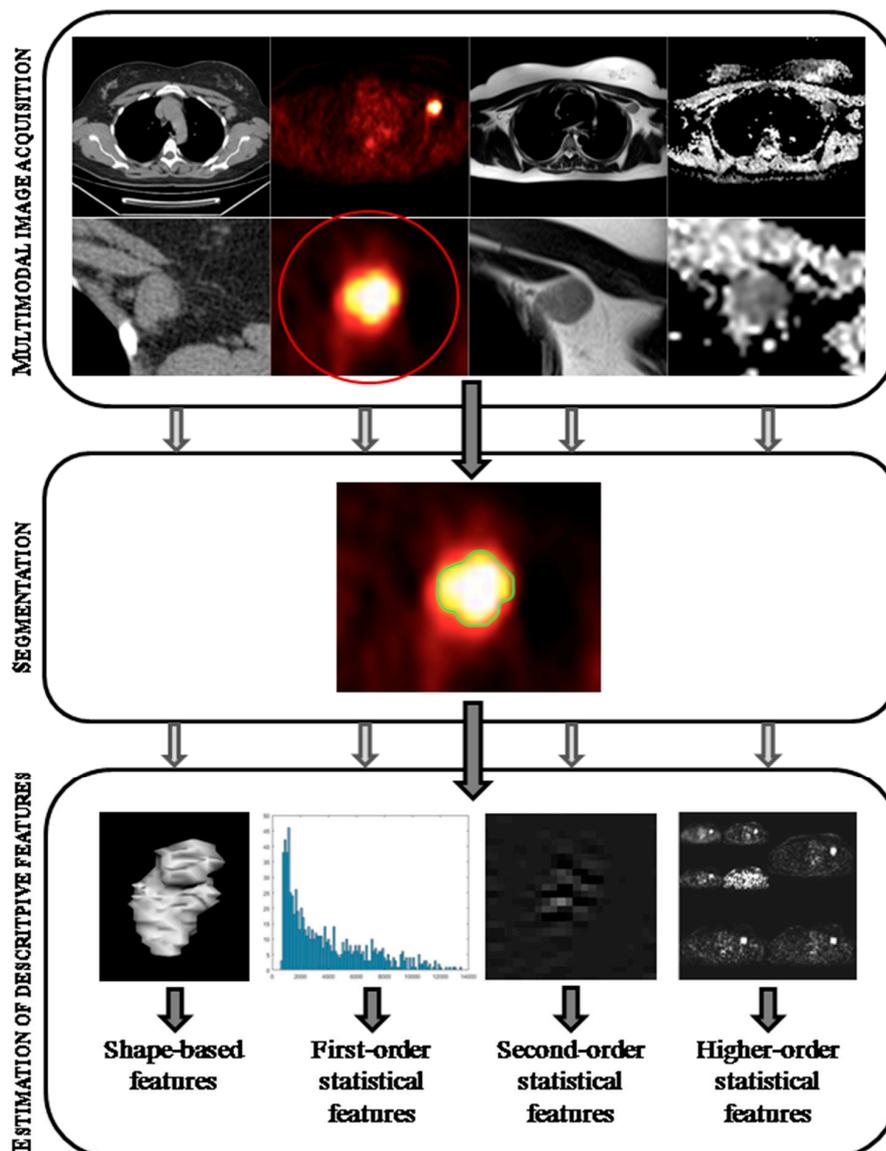


Figure 1. Schematic radiomic workflow. The process starts with the acquisition of common diagnostic images (Computed Tomography (CT), Positron Emission Tomography (PET), Magnetic Resonance (MR) see Section 2.1.1) and the identification of the lesions under investigation. The target regions are segmented (for the sake of simplicity, the process is shown for a single regions of interest (ROI) in the PET image only, as highlighted by the red circle) with the chosen approach (see Section 2.1.2). Finally, for each segmented region up to some hundreds of features, which are typically divided in shape-based, first-, second- and higher-order statistical features, can be computed (see Section 2.1.3).

Although the basic concepts underlying the radiogenomics are essentially intuitive, the definition of their scope is rather controversial [8,9]. Two main reasons of ambiguity can be identified. The first one arises from the prefix radio that may be interpreted as referring to radiation leading to the radiogenomics meant as radiation genomics. This should aim to develop an assay able to predict which cancer patients might develop toxicity as a result of radiotherapy treatment and to identify genes possessing single nucleotide polymorphisms (SNPs) as possible biomarkers of radiation-induced adverse effect. The second reason lies in the suffix omics: it implies the generation of complex high-dimensional mineable data from each biological and imaging sample [10]. This requirement is not completely fulfilled by those radiogenomic studies that simply conduct correlations between low-dimensional data of imaging and genomic data.

Here, we report the most recent advances in molecular profiling research and technology applied to cancer, and their integration with imaging features. In the following sections, radiomic and genomic techniques will be treated separately, and will be followed by a subsection regarding their integration and data analysis approaches. In the Discussion section, the radiogenomics approach will be outlined for specific neoplastic diseases, separately analyzed by specifying the applied radiomic and genomic methods, also considering the goal of each study. Finally, in the last section, critical issues and future perspectives will be put forward.

In conclusion, this review aims to critically illustrate the most relevant radiomic and genomic techniques applied in the radiogenomics field, and how this new methodological approach was used in predictive and correlation oncological studies.

2. Methodologies

2.1. Radiomics

Radiomics refers to the comprehensive quantification of tumour phenotypes by the extraction of a large number of quantitative features from medical images. This high-throughput extraction of quantitative imaging features is the result of a workflow that is composed of three main steps (Figure 1):

- Acquiring the images
- Segmenting the regions of interest (ROIs)
- Estimating descriptive features.

Each point involves well-established medical image acquisition and processing techniques, but the integration of these procedures into a reliable and reproducible pipeline for radiogenomic analyses deals with complex and challenging issues in each step. In what follows, these steps will be separately described and discussed.

2.1.1. Image Acquisition

A great advantage of radiomic analyses is their feasibility relative to conventional clinical scanners and imaging techniques. Indeed, the first step of the radiomic pipeline involves the acquisition of images that are typically part of diagnostic or treatment planning protocols for oncological patients [11], without the addition of examinations or extension of a scanning session, which could be unpleasant for the subject. However, when images are analyzed numerically to extract meaningful data, as is done in radiomic studies, variations in acquisition and image reconstruction parameters can introduce changes that are not necessarily due to underlying biological effects [5]. Consequently, the used acquisition protocols should be highly standardized in order to collect a dataset that is suitable for reliable analyses avoiding issues that can create difficulties in comparing results. Radiomics can be performed using tomographic images from Computed Tomography (CT), Magnetic Resonance (MR) imaging and Positron Emission Tomography (PET) [12] studies. Each modality has its own imaging issues, which

will be pointed out in the following subsections. Significant efforts are required to identify univocal acquisition and reconstruction protocols and to match them between scanners [13].

Computed Tomography (CT)

Relevant parameters have to be taken into account to obtain standardized protocols are related not only to image acquisition, but also to the reconstruction procedure, which delivers the image to the analysis step. Slice thickness, current (mA), and tube potential (kVp) are responsible for the photon statistics within a slice, therefore affecting the histogram and the noise level of an image. The axial field of view and the reconstruction matrix size change the voxel size within a slice, while the pitch, which is frequently optimized by each scan manufacturer, is crucial for the control of image noise. With respect to the reconstruction algorithm different approaches can lead to changes in Hounsfield units [13]. This may result in different quantitative features being extracted from the same acquisition protocol, with obvious negative effects on the reproducibility of radiomic analyses.

Positron Emission Tomography (PET)

Many issues related to the standardization of PET acquisition can be attributed to the calibration and quality controls are necessary to perform a quantitative examination [14]. With respect to CT, the scan and reconstruction parameters have to be standardized. For example, differences in grid size and post-reconstruction filter width imply large variations in the extracted features [15]. Additionally, the partial volume effect (PVE), which is mainly due to limited spatial resolutions and a relatively high noise contributions from PET systems, may quantitatively affect the voxel values and, consequently, the extracted features, especially in the case of small patterns that are comparable to the spatial resolution [16]. Consequently, there is a great interest in PVE correction methods that improve the reliability of quantitative index extraction [17]. Moreover, when dealing with PET quantitative acquisition, a precise patient protocol also has to be taken into account, with respect to dose calibration, blood glucose level, and uptake period [18].

Magnetic Resonance (MR)

Due to its complexity, MR can potentially play an important role in radiomic studies. MR images provide multi-parametric information, according to the used contrast weighting, which can be useful in categorizing a tumour. Ideally, MR images should all have, at least, the same field of view and acquisition matrix, field strength, and slice thickness, as these parameters strongly affect the Signal-to-Noise Ratio (SNR). Moreover, each acquisition technique requires specific expedients. For example, the commonly-used diffusion-weighted imaging (DWI) technique is strongly dependent on the k-space trajectory, gradient strength, and b-values, while dynamic contrast-enhanced (DCE) MR, and dynamic susceptibility contrast (DSC) MR depend on the contrast agent dose, administration procedure, and pulse sequence [13].

2.1.2. Region of Interest Segmentation

This step requires a preliminary identification of target regions of prognostic value. Radiomics is based on the concept that restricting the features to mine at the input point may be inefficient, while it is worth capturing as much data as possible, and identifying the features with the highest prognostic values, only at the end-point of the process using database mining [5]. Consequently, sub volumes of interest within the lesions, representing physiologically-distinct volumes (habitats), can be captured and added to the analyses [19,20].

Once the volumes of interest have been identified, the segmentation strategy has to be chosen. This point is very critical as the resulting feature values depend on the adopted strategy and the segmentation accuracy. In radiomic studies, the preferred segmentation method should be time efficient and should also provide accurate and reproducible boundaries. Usually, manual segmentation by expert readers is considered the gold standard, but it is a time-consuming process with high

inter-operator variability [21,22]. On the other hand, many automated methods have been developed across various image modalities and anatomical regions. They are completely reliable and reproducible in ideal conditions (such as normal structures and the absence of image artifacts), but they can fail in pathological situations, above all in complex cases of tumours with indistinct borders, because of inter- and intra-subject morphologic and contrast heterogeneity. Consequently, the best compromise has been identified in computer-aided detection systems [5,13,23] that work semi-automatically, with minimal user interaction (i.e., identification of seed points or manual correction). The use of semi-automated methods has also paved the way for three-dimensional (3D) segmentation. Volumetric segmentation allows a comprehensive view of the total tumour burden and volumetric assessment has shown good performance when related to different end points [24,25]. When comparing 3D with 2D approaches, it should be taken into account that tridimensional regions of interest, not only allow for a more complete description of the shape of the lesion [26], but also increase the number of points included in the statistical feature computation, leading to more reliable results, which do not suffer from sampling errors [27]. However, volumetric segmentation cannot be tackled with a standard, labor-intensive, manual approach. By reducing the manual workload, computer-aided approaches allow fast and reproducible 3D volumetric segmentations in large cohorts of patients, such as in radiomic studies [23].

2.1.3. Descriptive Features

Once each ROI has been segmented, radiomic analyses are based on the automated extraction of features that robustly and quantitatively describe the attribute and the complexity of individual ROIs. These quantitative image features should offer information on the tumour phenotype and microenvironment, which represent an evolution of the semantic features, which have been introduced in the clinical evaluation of several oncological lesions (BI-RADS for breast cancer [28], PI-RADS for prostate cancer [29], VASARI for glioblastoma [30]). Semantic features are qualitative extracted by a radiologist that follows a controlled lexicon to describe lesions and represents a coarse attempt to approach radiomics. Their quantitative translation with mathematically-extracted descriptors, which go beyond of the expert eye of radiologist, can complement and overcome them with an operator-independent and high throughput approach.

Typically, extracted features can be divided into shape-based, as well as first-, second-, and higher-order statistical outputs.

Shape-Based Features

Shape-based features rely heavily on the segmentation approach used, and capture numeric information regarding geometric characteristics, such as size, shape, and spiculation. Some intuitive examples are:

- Volume:

$$V = N \cdot vs \quad (1)$$

where N is the number of voxels within a segmented volume of interest, and vs is the voxel size of the acquisition.

- Surface area:

$$A = \sum_{i=1}^{NT} \frac{1}{2} |a_i \vec{b}_i \times a_i \vec{c}_i| \quad (2)$$

NT is the number of triangles obtained from the triangulation of a tumour surface; a_i , b_i , and c_i are the vertices of the i -th triangle. The surface area, together with the volume, and eventually, the maximum diameter, provide information on the size of a lesion.

- Compactness:

$$c = 36\pi \frac{V^2}{A^3} \quad (3)$$

This factor measures how much a lesion is different from a sphere, indicating, consequently, its irregularity.

First-Order Statistics

These features describe the distribution of voxel values without concern for their spatial relationships. They are generally histogram-based and can be used to quantify phenotypic traits [31]. Some examples are:

- Mean: shows the average intensity value and is given by:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X(i) \quad (4)$$

$X(i)$ is the gray value of the i -th voxels within a region of interest. Other estimates of the central tendency, used in descriptive statistics, can be computed, such as the mode and median.

- Standard deviation: indicates how widely intensity values vary, and is computed as:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X(i) - \bar{X})^2} \quad (5)$$

Other measures of histogram dispersions are the variance and the mean absolute deviation. The variability within a volume can also be indicated by common statistics, such as minimum, maximum and range values.

- Entropy: a statistical measure of randomness within a data sample, given by:

$$entropy = - \sum_{i=1}^{N_i} P(i) \log_2 P(i) \quad (6)$$

where P is the first-order histogram of the volume of interest, computed on N_i bins. Additionally, uniformity and energy can be used to measure the randomness of a volume histogram.

- Skewness: a parameter that describes the asymmetry of a histogram around the mean, calculated as:

$$skew = \frac{\sum_{i=1}^N (X(i) - \bar{X})^3}{N\sigma^3} \quad (7)$$

- Kurtosis: a parameter that depicts the degree of peakedness (broad or narrow) of a histogram and is given by:

$$kurt = \frac{\sum_{i=1}^N (X(i) - \bar{X})^4}{N\sigma^4} \quad (8)$$

Second-Order Statistics

Second-order statistical descriptors, typically referred to as “texture” features, describe spatial relationships between voxels with similar gray levels within a lesion. They provide a measure of intralesional heterogeneity [32]. Typically-used techniques are the Gray Level Co-occurrence Matrix (GLCM) [33], also known as Haralick features, the Gray Level Run-Length Matrix (GLRLM) [34], and the Gray Tone Difference Matrix (GTDM) [35].

- Gray Level Co-occurrence Matrix: These matrices determine how often a pixel of intensity i finds itself within a certain relationship to another pixel of intensity j . A GLCM is a joint probability function, defined as $P(i,j;d,a)$, where the elements (i,j) represent the number of times that intensity levels i and j occur in two voxels separated by distance d in the direction a . The matrix size

depends on the intensity levels within a segmented lesion and the number of matrices on the chosen d and a . For each matrix, several features can be extracted, and the final value for each d considered is obtained as the mean of the feature over the directions. Examples of characteristics that are mineable from each matrix are: Mean, standard deviation, and entropy for the joint and marginal probabilities, autocorrelation, cluster prominence, cluster shade and tendency, contrast, correlation, difference entropy, dissimilarity, energy, homogeneity, etc. [36].

- **Gray Level Run-Length Matrix:** A gray level run is the number of consecutive pixels having the same grey levels. In a GLRLM, defined as $p(i,j;a)$, the row indices represent the discretized gray values and the column indices are the number of consecutive occurrences of the i -th gray value in direction a . The matrix size, consequently, depends on the number of gray values in a lesion (number of rows) and the maximum run length (number of columns). A GLRLM can be obtained for each a , and the textural features can be obtained as the mean over the directions of the values extracted from each matrix. Examples of mineable features are: Short and long run emphasis, gray level non-uniformity, run-length non-uniformity, run percentage, low and high gray level run emphasis, etc. [30].
- **Gray Tone Difference Matrix:** A column matrix, in which elements $s(i)$ are the sum over the set of pixels having gray tone i , of the difference between the voxels of the set and the mean value, computed over the corresponding neighborhood. Consequently, the matrix depends on the size of the neighborhood. From GTDM, several features can be computed: Coarseness, contrast, busyness, complexity, and strength.

Higher-Order Statistics

Higher-order statistics impose filter grids on an image to extract repetitive or non-repetitive patterns. From filtered images, first- or second-order features are computed. Filters that are most often used in radiomic studies are:

- **Laplacian of Gaussian filter [32]:** This allows the highlighting of structures at a particular scale, corresponding to the width of a filter. Consequently, increasingly coarse texture patterns can be extracted from an image and analyzed using second-order statistics.
- **Gabor filters:** These allow for edge detection in different directions and widths [37]. For each filtered image, the Gabor energy feature can be extracted as a sum of the square intensity over all tumour pixels.
- **Wavelet transform:** This decouples textural information by decomposing the input image into low- and high-frequency coefficients without losing spatial localization. In particular, high-frequency coefficients also contain information on texture directionality. If an undecimated scheme is chosen, lesion segmentation, identified in the original image, can be used for computation of the first-order statistics and textural features from the wavelet coefficients.
- **Fractal dimensions:** These are estimates of object complexity. Fractal dimensions describe the relationship between the change in a measuring scale and the measurement at that scale [31], and can be calculated using a 3D box-counting algorithm [15]. Successively, values, such as mean and standard deviation, can be extracted.

As a result of features extraction, up to hundreds of features will be delivered to the statistical analysis, for only the radiomic part.

2.2. Genomics

Genomics is an interdisciplinary field of science focusing on genomes, which is actually performed using a combination of high-throughput molecular biology technologies with complex computing and math techniques (bioinformatics analysis). Nowadays, technological improvements have provided the possibility of determining the expression of thousands of genes involved in different fields of human

health and pathologies. Generally, two technologies are critical for genomics analyses: (1) microarray; and (2) next-generation sequencing (NGS) including DNA and RNA sequencing (Figure 2).

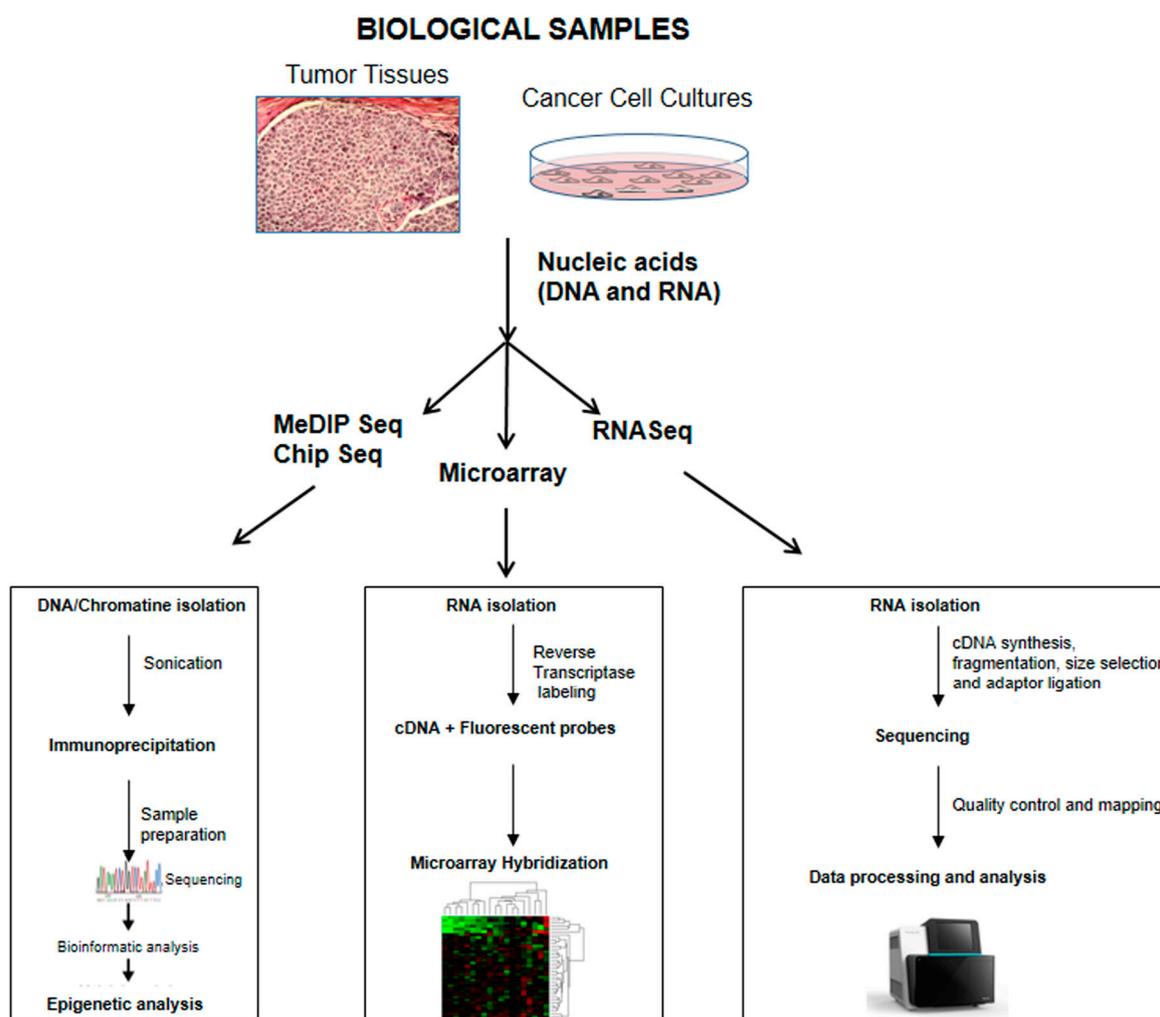


Figure 2. Schematic workflow of the most used genomic approaches that could be applied in the field of radiogenomics. Methylated DNA immune precipitation sequencing (MeDIP Seq) and Chip sequencing (Chip Seq) provide information about DNA methylation, DNA/protein interactions, and histone modification. Microarray is a technique used to measure the expression levels of large numbers of genes. RNA sequencing (RNA Seq) allows performing an in-depth analysis of the transcriptome with the identification of novel transcripts, alternative splicing allele specific expression, gene fusions, and genetic variations. Moreover, RNA Seq can give information about the transcriptome dynamics such as RNA editing, small insertions/deletions, exon connections, non-coding RNAs, and small RNAs.

2.2.1. Microarray

Microarray technology allows the understanding of complex functional mechanisms that are involved in physiological and pathological cellular processes. Indeed, DNA microarrays provide a rapid and accurate analysis of global gene expression, in contrast to previous methods for quantifying mRNAs, such as Northern blotting or quantitative PCR, that are able to measure a few genes at a time. DNA microarrays method, individual PCR-amplified double-stranded cDNA fragments (about 500 base pairs) are spotted onto microscopic glass slides; typically this technique exploits a spotted two-color hybridization to visualize and measure gene expression levels when comparing gene expression profiles from multiple samples [38,39]. In case of comparisons across a large number of

samples, a common reference sample is usually used as standard in all experiments [40–42]. The result of a DNA microarray is like a snapshot of actively expressed genes and transcripts (transcriptome) at a given point in time. However, although microarrays are used for identifying the expression of known genes and transcripts, this method fails to detect unidentified genes or transcripts. The solution to this issue can be circumvented by the use of in-depth sequencing analyses, such as Next-generation sequencing (NGS) analyses, discussed below. DNA microarrays are also exploited in the field of epigenetics, which refers to changes in gene expression occurring without alterations to the DNA sequence. It is known that many neoplastic transformation events are connected to deregulated epigenetic machinery, such as DNA methylation of the CpG island within promoter regions and the covalent modification of histone proteins, such as acetylation, phosphorylation, methylation, and ubiquitination, are able to regulate gene activity [43]. For large sample size genome-wide DNA methylation studies, the most used platform is the Infinium HumanMethylation450 BeadChip kit (Illumina, San Diego, CA, USA). This technology is able to obtain a rapid quantitative DNA methylation analysis of more than 485,000 CpG dinucleotides (methylation sites) located across the genome [44] (Figure 2). As reported in the Discussion section, most radiogenomics studies, on a variety of human cancers, used microarray technology. In these studies, genomic data were compared using different radiomic approaches such as MRI, CT, and PET. The aim of these studies was to identify imaging and genomic features that are able to ameliorate the diagnosis and prognosis of cancer patients.

2.2.2. Next-Generation Sequencing

Next-generation sequencing (NGS), also known as high-throughput sequencing, comprises a number of different modern technologies that are able to provide more accurate information than the previously-used Sanger method, revolutionizing the study of genomes and representing new opportunities for clinical applications [45,46]. The NGS-era started in 2004 with the introduction of massively parallel sequencing platforms [47] (e.g., the Applied Biosystems and the Helicos BioScience HeliScope) able to produce 400 million of 25–35 bp reads [48]. Although there are different technical details, a single DNA molecule in each sample is, first, fragmented, amplified, and then sequenced. A common feature among NGS technologies is that the template is attached to a solid surface or support, which enables, simultaneously, thousands, and up to billions of sequencing reactions. The massive amount of data generated is attractive, and the elimination of the cloning step for the DNA fragments to be sequenced is the greatest benefit of these new technologies. The broadest application of NGS may be the resequencing of different genomes and, in particular, human genomes for an in-depth understanding of genetic differences in health and disease. Furthermore, these platforms have been used in many genomic applications, leading to a wide range of so-called “Seq” protocols, such as RNA-Seq for transcriptomics, Chip-Seq for DNA-protein interaction analyses, and methyl-Seq for epigenetic mark profiling. Several programs can be used for data handling and visualization, quality assessment, interpretation, and statistical analyses [49–51] (Figure 2).

RNA-Sequencing

Particular interest in oncology is focused on the field of transcriptomics for the identification and quantification of the RNA in cells, tissues, or biological fluids, representing a powerful tool for the assessment of specific biological activities [52]. In a single RNA-Seq experiment, it is possible to investigate, not only gene expression, but also alternative splicing [53], novel transcripts [54,55], allele specific expression [56], gene fusions [57], and genetic variations [58,59]. Moreover, RNA-Seq can provide more interesting information regarding transcriptome dynamics, such as RNA editing, small insertions/deletions, exon connections, non-coding RNAs, and small RNAs [60]. Today, there are three widely-accepted, commercially-available NGS devices for RNA-Seq: 454 GS FLX (up to 400 bp) (Roche, Basel, Switzerland), Genome Analyzer II, with paired-end reads up to 100 bp (Illumina, San Diego, CA, USA), and SOLiD (up to 35–50 bp) (Applied Biosystems, Foster City, CA, USA) [61,62]. Although each platform works differently, they are all based on similar principles: Shearing target

nucleic acids into small pieces, binding individual molecules to a solid surface, amplifying each molecule into a cluster, copying one base at a time, and detecting different signals for each nucleotide base. The majority of the platforms only allow for the sequencing of DNA molecules. Therefore, RNA molecules are, first, reverse transcribed into cDNA. Once reverse-transcription is complete, the RNA molecule is removed. At the end of the entire process, the result is a sequence of images, where each lighted spot corresponds to a cluster and the color of each cluster represents a different base type [51].

The first step of RNA-Seq data analysis is the quality control of the raw reads, particularly the determination of sequence quality, GC content, overrepresented k-mers, and duplicated reads, in order to detect sequencing errors, PCR artifacts, or contamination [63]. An important indicator of sequencing quality/accuracy and absence of contaminating DNA is represented by the percentage of mapped reads [64,65]. After quality control, NGS-data analysis is performed by mapping the sequence reads. Indeed, reads are aligned to a reference genome, or to reference transcripts, or assembled *de novo* without a referenced genomic sequence to produce a genome-scale transcription map, consisting of the transcriptional structure and or the expression level for each gene. When a reference genome is available, RNA-Seq analysis involves the mapping of the reads onto the reference genome or transcriptome, even with the limitation of discovery new transcripts. If an organism does not have a sequenced genome, a *de novo* assembling approach to produce a genome-scale transcriptional map is necessary [66–69]. For the identification of novel transcripts, several software packages and algorithms are used to assess splice junctions and transcription start and end sites [70–75]. Gene-finding prediction tools, such as Augustus [76], can exploit RNA-Seq data to better annotate protein-coding transcripts [77].

The second phase of RNA-Seq analysis provides the quantification of transcript expressions using programs such as HTSeq-count [78] or feature Counts [79], based on the aggregation of the raw counts of mapped reads. Quantitative gene expression data from RNA-Seq have been shown to be comparable to those of microarrays, but with a better dynamic range and lower detection limit for low-expressed transcripts [59] (Figure 2).

RNA-Seq is also used to study the biological role and signature of small RNAs (sRNAs). Although sRNA-Seq libraries are rarely sequenced as deeply as classical RNA-Seq libraries, and bioinformatics analysis is different from standard RNA-Seq protocols, obtained sRNA reads are aligned to a genome or transcriptome reference using bioinformatic tools, such as Bowtie2 [80], STAR [81], or Burrows-Wheeler Aligner (BWA) [82].

Innovations in RNA-Seq have made quantitative transcriptome analysis of a single cell possible, even when RNA-Seq is performed on a large number of cells in the same run [83]. Furthermore, methods that integrate DNA whole exome sequencing (DNA-WES), or Chip-Seq, with RNA-Seq have allowed increased mutation detection performance [49,84,85]. Finally, an *in situ* method of RNA-Seq has also been developed for preserved tissue sections or cell samples [86].

RNA-Seq offers several advantages compared with other transcriptomics methods [59,87], providing high-throughput solutions for the construction of single-base resolution expression profiles with low background noise and a low amount of required starting RNA. Furthermore, it can generate millions of reads in a single run. Nevertheless, few studies used the RNA-Seq technology in combination with radiomic technologies (MRI) to address clinical issue, are reported in Table 1. Indeed, although RNA-Seq provides results that are superior to microarray analysis, in terms of sensitivity, specificity, and abundance estimation, microarrays are still used more than RNA-Seq. This is probably due to costs, run time, and the large volume of data, that make it necessary to dedicate platforms to data storage (Big Data). In light of this, although RNA-Seq is promising, technological improvements for reducing costs, improving data processing/storage, and gold standards for analyses are necessary to best use this powerful platform in research laboratories and clinics.

2.2.3. Immunohistochemistry

The utility of the immunohistochemistry (IHC) technique for the improvement of microscopic diagnosis of neoplasia is known. This method aims to characterize cellular or tissue constituents for example by identifying therapeutic targets (markers) for cancer by taking advantage of antigen-antibodies interactions. This method expects different steps: Deparaffinization of tissue sections, quenching of endogenous enzymes to avoid false positive results, antigen retrieval, blocking of nonspecific binding sites, binding primary antibodies, binding with biotinylated secondary antibodies, detection by using one or several methods as peroxidases- antiperoxidase, avidin biotin conjugates, peroxidase complexes, or by using a polymer-labelling two-step method, addition of chromogen substrate, counterstaining, dehydrating, and cover slipping the slide. Advantages of these methods include protein localization and distribution, applicable for different sizes of tissue biopsies, and fixed tissues and validation of other high-throughput studies (DNA microarray). Nevertheless, these methods show some limited ability regarding the detection of protein modifications and quantifications [88].

2.3. Radiogenomic Data Analysis

Once both radiomic and genomic features are extracted for each subject, radiogenomic analysis are performed on a population dataset, including the clinical outcome related to each sample. A prediction study is aimed to forecast either an overall or progression-free survival or response to a particular treatment [89]. In addition, prediction studies are needed to recognize a class of disease or to predict genotype from imaging phenotype and vice-versa [90]. A simple radiogenomic correlation study, by contrast, can be performed investigating the mutual statistical relationship between radiomic and genomic features for a given disease, regardless the clinical outcome [91].

Considering the very large dimensionality of the feature space, especially when dealing with multimodal imaging, advanced algorithms that rank features by their importance for a given disease outcome are often essential to reduce over fitting, increase reliability, and address the curse of dimensionality [92]. The latter, which refers to a crucial aspect in radiogenomic data analysis, can be roughly summarized as the requirement of higher sample sizes as the number of the features increase.

Feature selection can be performed as a preparatory step (filtering approach) or can be embedded within the classification or regression procedures, identifying a sparse set of features that can be used to train highly-accurate predictors of individualized outcomes. In References [93,94] features were ranked and selected on a test-retest reproducibility basis before proceeding with a correlation analysis; redundancy was avoided by first eliminating highly-correlated features [95]; see Reference [96] for a comprehensive description and evaluation of the most popular feature filtering methods. As an instance of embedded feature selection, the work in Reference [97] employed a LASSO regularization [98] within a logistic regression to assess the predictive power of each feature.

Since the domain of each feature can be different, data normalization strategies should be considered on the basis of statistical analysis methods; a well-established choice is the use of a z-score transformation [96], but, as previously demonstrated [36], radiomic data can be employed without any normalization processing.

Once each sample is arranged into a suitable feature space, data can be classified in a supervised or unsupervised manner. Unsupervised methods, such as clustering and principal component analysis, are data-driven methods aimed at discovering clusters, representing different radiogenomic phenotypes, within the feature space [99]. Supervised methods, such as machine learning approaches, involves a prior labelling of data containing the clinical outcome (e.g., overall survive/not survive; responder/non-responder) and a training step to infer a model underlying the data, which can be used for predicting the labels of unseen observations. Classifiers, such as support vector machines (SVM), span different areas of computer and information science; see Reference [96] for a description of well-established methods employed in radiogenomics. Alongside typical machine learning methods, the family of deep learning algorithms is gaining increasing interest as robust classifiers for the

radiogenomic domain [100,101]. It is worth mentioning that machine learning and correlation analyses do not solve the problem of causal inference in observational data sets [4]. Therefore, conventional radiogenomic methods carry out a correlative association of radiomic signatures with gene signatures. Alternatively, a recent study [102] proposed a preclinical radiogenomic methodology to demonstrate that image features are causally related to genetic factors.

Figure 3 shows the whole radiogenomics pipeline, in light of methodological issues described throughout this section.

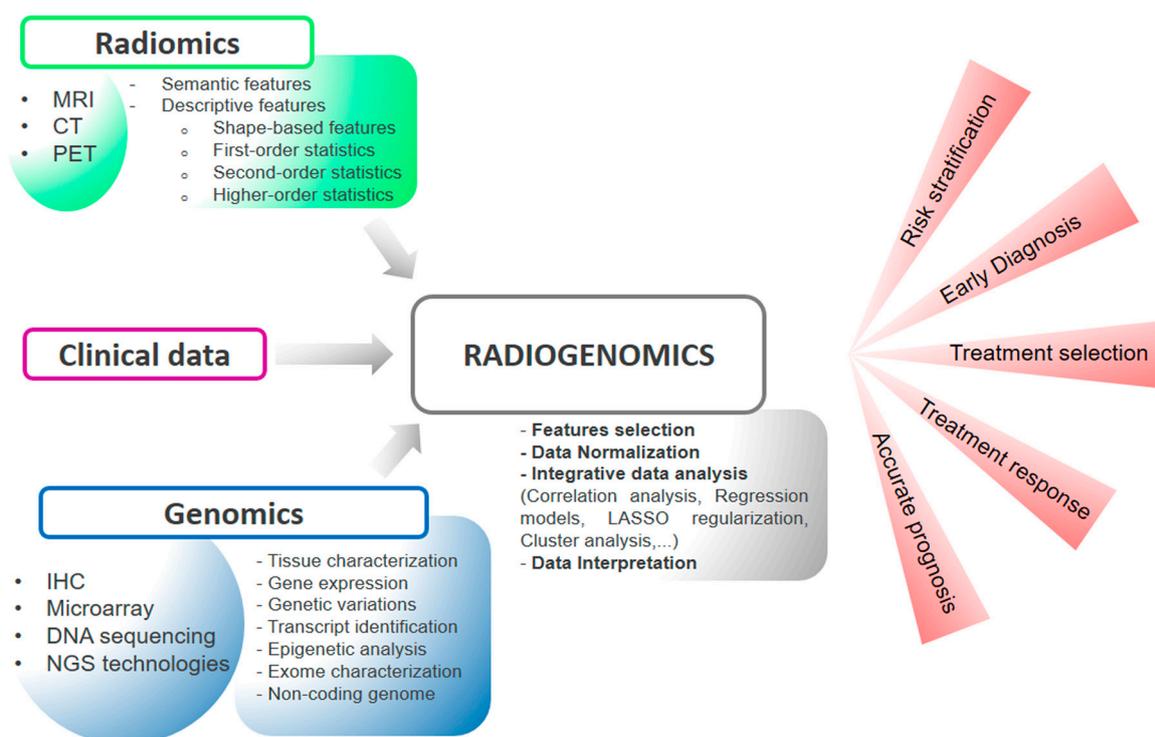


Figure 3. The figure shows a general workflow for radiogenomic study. The first step includes data acquisition (clinical information, imaging and genomic data). Subsequently, data are normalized and underwent an integrative analysis to characterize each radiomic feature and identify specific underlying molecular functions. The overall flow, here schematically depicted, could represent a novel integrated approach for cancer diagnosis and prognosis.

3. Discussion

In the last five years there has been a significant increase of studies that use the radiogenomic approach to obtain clinically-significant information to be translated into the clinical practice. As outlined in Figure 3, there are multiplicity of methods involved in radiogenomics, as well as wide variety of the possible outcome of a radiogenomic study.

Table 1 summarizes most of the relevant oncological studies in radiogenomics, crossing imaging and biological data. We searched multiple electronic databases for original research studies including the following words: “radiogenomics and breast cancer”, “radiogenomics and glioblastoma”, “radiogenomics and lung cancer”, “radiogenomics and kidney cancer”, “radiogenomics and renal cancer”, “radiogenomics and hepatic carcinoma” “radiogenomics and prostate cancer”. We excluded studies that associated the term radiogenomics to the radiation therapy response. The studies are classified according to the neoplastic disease and for each of them the methodological approach, as introduced in the previous sections, and the performed statistical analysis are pointed. In the next sections, the main results of the application of the radiogenomic approach as well as the techniques employed will be deeply discussed for each oncological disease.

Table 1. Radiogenomic studies published in oncology.

Tumour Type	Rationale of the Study	Number of Sample	Imaging Data	Imaging Features	Segmentation	Genomic Features	Statistical Analysis	Ref.
BREAST CANCER (BC)	Correlation	275	MRI (T1WI, T2WI, DCE)	Shape-based features, second- and higher-order statistics, kinetic parameters	Semi-automated	IHC	Binary multivariate logistic regression model and univariate models	[103]
	Prediction	91	MRI (DCE)	Shape-based features, second-order statistics, kinetic parameters	Semi-automated	RNA Seq Microarray (TCGA)	Logistic regression with LASSO regularization and ROC analysis	[97]
	Correlation	48	MRI (T1WI, T2WI, DCE)	Shape-based features, second-order statistics, kinetic parameters	Semi-automated	IHC	Multivariate logistic regression models	[104]
	Correlation	221	MRI (T1WI, T2WI, DCE)	Semantic features	Manual	IHC	Wilcoxon test and Fisher’s tests	[105]
	Correlation Prediction	95	MRI (T1WI, T2WI, DCE)	Shape-based features, first- and second-order statistics	Manual	IHC Microarray	Multiple linear regression analysis and Spearman’s rank correlation	[106]
	Correlation	178	MRI (T1WI, T2WI, DCE)	Shape-based features, first- and second-order statistics	Manual	IHC	Multiclass support vector machines with the a leave-one-out cross-validation approach	[107]
	Correlation	176	MRI (T2WI DCE)	Semantic features	Manual	IHC	Chi-squared and Fisher’s tests	[108]
	Correlation	353	MRI (T1WI, DCE)	Semantic features	Manual	Microarray	Spearman rank-correlation	[109]
	Correlation	109	MRI (T1WI, DCE)	Shape-based features, first- and second-order statistics, kinetic parameters	Automated	RNA Seq	Cox regression analysis	[110]
	Correlation	92	MRI (T2WI, DWI, DCE)	Semantic features, ADC	Manual	IHC	Mann–Whitney U and Kruskal–Wallis H tests	[111]
	Correlation	115	MRI (T1WI, T2WI, DWI, DCE)	ADC	Manual	IHC	Mann–Whitney U and Kruskal–Wallis H tests	[112]
	Prediction	50	MRI (T1WI, T2WI, DCE)	Semantic features	Manual	IHC	Student’s unpaired <i>t</i> -test, one-way ANOVA, Chi-squared and Fisher’s test	[113]
	Correlation	282	MRI (T1WI, T2WI, DCE)	Shape-based features	Manual	IHC	Multiple linear regression analysis	[114]
	Prediction	96	MRI (T1WI, T2WI, DWI, DCE)	Shape-based features, ADC	Manual	IHC	Multivariate logistic regression analysis	[115]
	Prediction	214	MRI (T1WI, T2WI, DWI, DCE) PET/CT	ADC, SUV	Manual	IHC	Mann–Whitney <i>U</i> and Kruskal–Wallis <i>H</i> tests	[116]
	Correlation	103	PET/CT	SUV	Manual	IHC	Chi-squared test, Fisher’s and Wilcoxon tests	[117]
	Correlation	552	PET/CT	SUV	Manual	IHC	Univariate and multiple linear regression analysis	[118]

Table 1. Cont.

Tumour Type	Rationale of the Study	Number of Sample	Imaging Data	Imaging Features	Segmentation	Genomic Features	Statistical Analysis	Ref.
BREAST CANCER (BC)	Correlation	82	PET/CT	SUV	Manual	IHC	Chi-squared test, Fisher’s and Mann Whitney tests	[119]
	Correlation	91	MRI (DCE)	Shape-based features, second-order statistics, kinetic parameters	Semi-automated	Microarray (TCGA)	Regression and clustering analysis	[91]
	Correlation	228	MRI (T2WI, DCE)	Kinetic parameters	Semi-automated	IHC	Kruskal–Wallis <i>H</i> test	[120]
	Prediction	36	MRI (DCE)	Kinetic parameters	Manual	IHC Microarray	Wilcoxon test, Spearman’s rank correlation, and Kruskal–Wallis <i>H</i> test	[121]
	Correlation	36	PET	SUV	Manual	IHC Microarray	Two-way unsupervised hierarchic clustering and Spearman’s rank correlation	[122]
	Correlation	18	PET	SUV	Manual	Microarray	Rank-rank hypergeometric overlap	[123]
GLIOBLASTOMA (GBM)	Correlation Prediction	25	MRI	Semantic features	Manual	Microarray	Correlation analysis	[124]
	Correlation	78	MRI-FLAIR, T1-c	Size, volume	Automated	TCGA	Pathways genomic analysis	[125]
	Correlation	23	MRI (T1-c, DSC)	Semantic features	Manual	Microarray (GSEA)	Correlation analysis	[126]
	Correlation Prediction	76	MRI (TCIA) MRI (T1-c, FLAIR)	Semantic features	Semi-automated	Microarray (TCGA)	Student’s <i>t</i> -test ,ROC AUC analysis	[127]
	Correlation Prediction	92	MRI (TCIA)	Semantic features	Manual	Microarray (TCGA)	Hierarchical clustering and survival analysis	[128]
	Correlation	48	MRI anatomical	Second-order statistics	NA	CGH array exome sequencing	Multivariate predictive decision-tree models	[129]
	Correlation Prediction	55	MRI (TCIA)	Semantic features	Manual	Microarray (TCGA)	Cox proportional hazards modeling and correlation analysis	[130]
	Correlation	21	MRI (DSC)	Mean values	Manual	Microarray	Cox regression analysis	[131]
	Correlation	152	MRI (DWI,DSC, SWI,T1WI,T2W2)	First-order statistics, Semantic features	Manual	Microarray	Hierarchical clustering	[132]
	Correlation	13	MRI (DWI, DSC)	Mean values	Manual	Microarray	Correlation analysis	[133]
	Correlation Prediction	52	MRI (T1-c,DSC)	Clinical scores	NA	Microarray	Univariate Cox proportional hazard models	[134]
	Prediction	78	MRI (T1-c, DSC)	Semantic features	Manual	Microarray (TCGA)	Proportional Hazards Model	[135]
	Prediction	71	MRI	Clinical scores	NA	Microarray	Multivariate Cox proportional hazard models	[136]
Prediction	104	MRI (T1, T2 CE)	Semantic features	Manual	Microarray (TCGA)	Univariate proportional hazards regression	[137]	

Table 1. Cont.

Tumour Type	Rationale of the Study	Number of Sample	Imaging Data	Imaging Features	Segmentation	Genomic Features	Statistical Analysis	Ref.
GLIOBLASTOMA (GBM)	Correlation	18	perfusion CT	Perfusion parameters	Manual	Microarray (TGCA)	Correlation analysis	[138]
	Correlation	46	MRI (DCE, FLAIR)	Semantic features	Manual	Microarray	Kruskal – Wallis H test	[139]
	Prediction	68	MRI (DCE, DWI, anatomy)	Semantic features	Manual	Microarray (TGCA)	Univariate Cox Regression models	[140]
	Correlation	27	MRS	Metabolite concentration	Manual	IHC, PCR	Correlation analysis	[141]
	Correlation	26	MRI (DCE, DWI, DSC, MRS)	Semantic features	Manual	IHC	Correlation analysis	[142]
	Prediction	108	MRI (DCE, DWI)	Semantic features	Manual	Microarray (TGCA)	NA	[143]
LUNG	Prediction	186	CT	Semantic features	Manual	PCR	Univariate analysis and multivariate decision tree models	[144]
	Prediction	138	PET/CT	Shape-based feature, second-order statistics, semantic features	Manual	Microarray	Generalized linear regression with LASSO regularization	[145]
	Correlation Prediction	355	PET/CT	SUV	Manual	Microarray	Student’s t-test, Wilcoxon test, Chi-squared and Fisher’s test	[146]
	Prediction	422	CT	Shape-based features, first-, second- and higher-order statistics	Manual	Microarray	Intraclass correlation coefficient, Friedman test	[36]
KIDNEY	Prediction	70	CT	First-order statistics, semantic features	Manual	Microarray	Multivariate linear regression	[99]
	Correlation	233	CT	Shape-based features, first-order statistics, Semantic features	Manual	DNA-Seq (TCGA)	Fisher’s tests	[147]
	Correlation	103	CT and MRI	Shape-based features, first-order statistics, Semantic features	Manual	Microarray (TCGA)	Pearson’s test and Mann–Whitney <i>U</i> test	[148]
	Prediction	58	CT (TCIA)	Shape-based features, first- and second-order statistics	Manual	Microarray (TCGA)	Support vector machine classifier	[149]
LIVER (HCC)	Correlation	30	DCE-CT	Semantic features	NA	Microarray	Correlation analysis	[94]
	Correlation Prediction	47	three-phase contrast enhanced CT	Semantic features	NA	Microarray	Bayesian models	[150]
	Correlation	77	Liver-specific contrast enhanced-MRI	Clinical scores	NA	IHC Microarray	Student’s <i>t</i> -test	[151]
PROSTATE	Correlation	45	MRI (T1WI, T2WI, DWI, DCE)	First-order statistics, kinetic parameters, ADC	Manual	IHC	Spearman’s rank correlation coefficient	[152]
	Prediction	17	MRI (T2WI, DWI, DCE)	First-order statistics, kinetic parameters, ADC	Semi-automated	Microarray	Pearson’s correlation, two-way hierarchical clustering	[153]

ADC: apparent diffusion coefficient; DCE: dynamic contrast-enhanced; DSC: dynamic susceptibility contrast; DWI: diffusion weighted imaging; IHC: immunohistochemistry; MRS: Magnetic Resonance Spectroscopy; NA: not applicable; T1WI: T1 weighed imaging; T2WI: T2 weighed imaging; T1-c: T1 weighed post contrast.

3.1. Radiogenomics in Breast Cancer

The election imaging method for Breast Cancer (BC) was MRI (78%) [91,97,103–116,120,121]. About 78% of the BC published papers were correlation studies [91,103–112,114,117–120,122,123] and the most used genomic data was IHC results (65%) [103–105,107,108,111–120]. However, although IHC is routinely applied in the clinical practice, it is not an “omic” approach, remaining the elective method to determine the BC molecular subtypes (Luminal A, Luminal B, basal-like, and triple-negative (TNBC)), useful for therapeutic indications [154].

The first study, dates back to 2009, in which authors found a significant correlation between high histological grade, unifocal lesion, mass lesion type, smooth mass margin, rim enhancement, persistent enhancement pattern, and very high intratumoural signal intensity on T2W MR images with the triple-negative subtype [108]. This finding was confirmed by Sung and colleagues [105]. Cipolla et al. and Molinari et al. reported a significant correlation between ADC values with tumour aggressiveness in terms of grade and the Ki67proliferation index [111,112]. In addition, the Ki67 index and ER score correlated with the roundness of tumour, although in an independent manner [114]. Only three studies correlated IHC results and PET/CT imaging features, reporting an association between FDG uptake and Ki67 [117,119], FDG uptake and tumour size in TNBC [117] and higher SUVmax with triple-negative and HER2-positive tumours [118]. Three works focused on prediction studies combining IHC with MRI [113,115,116] or with PET [116].

Few works focused on correlation (21%) and predictive (12.5%) studies performed by the use of genomic data (DNA microarray or RNA Seq) in combination with MRI [91,97,106,109,110,121] and PET features [122,123]. According to the primary goal of radiogenomics, as described in previous sections, these studies clearly fulfill the radiogenomics concept. Specifically, Yamamoto et al. evidenced that contrast-enhanced MRI systematically correlated with complex transcriptome profiles [109] paving the way for the generation of imaging biomarkers correlated with individual patient biology. Recently, Sutton et al. evidenced in patients with invasive ductal carcinoma that MR imaging-derived features were correlated with the Oncotype Dx RS scores based on the expression levels of 21 BC-associated genes [106]. The ODx score correlated with the magnitude of chemotherapy, as well as the 10 years risk of distant metastases, representing a possible prognostic and predictive surrogate genomic-based test. The usefulness of in-depth imaging analysis of BC vascularity by MRI, using the Dynamic Contrast-Enhanced (DCE) parameters, has been witnessed in a recent study by Mehta et al. [121]. The authors found that DCE MRI could provide criteria for patient stratification in the case of an anti-angiogenesis trial. Indeed, patients with a high Ktrans value at diagnosis disclosed a higher complexity in terms of proliferating cancer genes, as well as increased vasculature. Based on these findings the authors postulated that DCE-MRI could evidence, in a non-invasive way, candidates that could benefit from anti-angiogenic therapy. Future studies are needed to evaluate, in large BC patient cohort, the clinical usefulness of the MRI-derived imaging biomarkers. In this context it is important to consider a recent paper where, for 91 breast-invasive carcinomas, it was possible to integrate The Cancer Genome Atlas (TCGA) data with the MRI data from The Cancer Imaging Archive (TCIA) [91]. In this work quantitative MRI phenotypes, such as tumour size, shape, margin, and blood flow kinetics, were associated with their corresponding molecular profile, including DNA mutation, gene, miRNA, and protein expression, as well as copy number variation. Only two studies used RNA-Seq methods in combination with MRI features [97,110]. The first one, a predictive study, analyzed the genomic and radiomic data of 91 BC patients from TCGA and TCIA in order to predict clinical outcomes [97]. The work of Yamamoto, instead, was a correlation study which introduced, for the first time, the use of long non-coding RNA (lncRNA) as genomic data combined with MRI features. A radiogenomic analysis allowed identifying eight lncRNAs, three of which were unnamed. The pathway analysis showed that known lncRNAs molecules were involved in metastasis, cell cycle, cell death and survival, cellular development, and cellular growth and proliferation. In particular, higher levels of lncRNAs HOTAIR and LINC00511 were found in tumour samples of invasive ductal carcinoma than in normal breast tissue with a significant positive correlation with texture features [110].

3.2. Radiogenomics in Glioblastoma Multiforme

From Table 1, the wide use of semantic features in GBM appears obvious; these are preferred to descriptive features, more suitable for radiomic analyses, due to their high level of standardization. It is worth to mention that the majority of GBM radiogenomic study exploits the intrinsic multimodality of MR, considering both anatomical and functional features. Therefore, the radiogenomic approach has proven successful in determining the MRI phenotype associated with genetic alterations detected via DNA microarray analysis in GBM. The first study dates back to 2007. In this study, the authors evaluated Vascular Endothelial Growth Factor (VEGF) and related gene expression in 71 malignant GBM tissues in order to analyze their relationship with edema and survival. By using DNA microarray and MRI methods, they found that VEGF expression was predictive of survival with tumours with little or no edema in GBM-affected patients [136]. Subsequently, Diehn and colleagues combined neuroimaging (MRI) and DNA microarray analysis in a correlation/predictive study to create a multidimensional map of gene expression patterns, which provided clinically relevant insights into tumour biology. This study offered a potential strategy for noninvasively selecting patients who may be candidates for individualized therapies [124]. In the same year, in a similar study, the relationship between gene expression and MRI enhancement in GBM was evaluated. Results showed that gene expression was able to discriminate magnetic resonance imaging features in incompletely enhancing (IE) and completely enhancing (CE) tumours and to predict patient survival [134]. In order to increase the use of non-invasive imaging as an emerging field of treatment response and personalized medicine, over the years, other works have been published in the field of the radiogenomics for the study of brain tumours, through the analysis of gene expression and DNA copy number variations in tumour tissues. Most of these are correlation studies that were performed using DNA microarray and MRI imaging modalities [125,126,129,131–133,139,141]. On the other hand, a plethora of studies investigated the possible predictive value of the radiogenomics approach, showing promising results [127,128,130,135,137,140,143]. The only study that used the CT modality associated with a DNA microarray assay showed a correlation between anti- and pro-angiogenic genes with tumour perfusion parameters [138]. In addition, recently, Demerath and colleagues used MRI and the less innovative molecular technique of IHC, to focus their study on the use of mesoscopic measures as surrogate markers for specific gene expression patterns in GBM [142].

3.3. Radiogenomics in Lung Cancer

Most radiogenomics works applied to lung cancer are predictive study. One of the first works, dating back to 2012, was focused on CT and PET images features in patients with non-small cell lung cancer (NSCLC) [145]. The authors explored the relationship between differential genome-wide expression using DNA microarrays and different FDG uptake levels in NSCLC, finding a gene expression signature associated with prognostically-relevant FDG uptake features. In particular, they exploited public gene expression microarray data and 180 image features from CT and PET/CT, identifying 243 statistically-significant pairwise correlations between image features and metagenes to assess prognosis and therapeutic response [145]. Subsequently, in a large cohort of patients with diagnosed NSCLC, Nair and colleagues found that Nuclear factor- κ B (NF- κ B) IHC expression was related to tumour metabolism, measured using FDG-PET, and prognosis provided a methodology for studying tumour biology using computational approaches [146]. Four hundred and forty quantitative image features, describing tumour phenotype characteristics (tumour image intensity, shape, texture and multiscale wavelet), were defined, and Aerts et al. compared these radiomic signatures with gene-expression profiles using gene-set enrichment analysis (GSEA) [36]. They found that a large number of radiomic features had prognostic power in lung and head-and-neck cancer patients, many of which were not previously identified as significant. Indeed, this radiogenomic approach revealed that a prognostic radiomic signature, capturing intratumour heterogeneity, was associated with gene-expression patterns, providing an opportunity to improve decision-support in cancer treatment at low cost. Recently, whether epidermal growth factor receptor (EGFR) and KRAS mutation status

can be predicted using CT imaging data, as been investigated [144]. In this study, the authors found a statistically-significant model for predicting EGFR, but not for KRAS mutations, showing the potential of quantitative imaging to predict molecular properties in a non-invasive manner.

3.4. Radiogenomics in Kidney Cancer

Few works using radiogenomic models have been published in the field of kidney carcinoma since 2014. These preliminary studies revealed the association between CT features of clear-cell renal cell carcinoma (ccRCC) with somatic mutations in several genes detected by PCR amplification and DNA sequencing. These genes (*VHL*, *PBRMI*, *SETD2*, *KDM5C*, *BAP1* and *MUC4*) were known to be related to advanced grade, stage, and reduced survival prognosis [147,148]. Study published by Karlo et al. was upgraded by the work of Jamishidi et al., where the authors constructed a complex, multi-feature imaging predictor by using a multi-gene predictive gene expression signature and 28 CT images [148]. Through these discoveries, the authors identified imaging features that were potentially predictive of outcomes.

3.5. Radiogenomics in Prostate Cancer

Several studies have addressed the utility of MRI to determine the aggressiveness of prostate cancer using ADC, DCE, and Gleason score parameters. To our knowledge, the first studies in prostate cancer, investigating the associations between MRI parameters with the genomic markers, were published in 2016 [152,153]. McCann and coauthors studied the correlation between quantitative imaging features of multi-parametric MRI and phosphatase and tensin homolog (PTEN) protein expression by IHC analysis. They found a significant, although weak, associations between the extracellular space and plasma (K_{ep}) and the Gleason score with PTEN expression, claiming that the use of this model can improve the risk assessment of patients with prostate cancer. The study performed by Stoyanova et al. showed a correlation between radiogenomic parameters and prostate cancer, exploiting MRI-guided biopsy. They identified radiomic features in normal-appearing tissues associated with high-risk gene expression profiles, as well as radiomic biomarkers in cancer tissues associated with genes of adverse outcomes. Gene expression profiles of tissue specimens were determined using DNA microarray.

3.6. Radiogenomics in Liver Cancer

By using an integrated imaging-genomic approach, the first study on liver cancer dates back to 2007. The aim of that work was to determine whether conventional contrast-enhanced CT could be used in identifying imaging phenotypes associated with a doxorubicin drug response gene expression program in hepatocellular carcinoma (HCC) [94]. They found that tumour margins in arterial phase images showed significant correlation with (i) the doxorubicin-response gene expression program; (ii) HCC venous invasion; and (iii) tumour stage. Furthermore, tumours with higher tumour margin scores were associated with the doxorubicin resistance transcriptional program, and had a greater prevalence of venous invasion a worse stage. In the same year, a study was published showing that 28 dynamic imaging traits in CT correlated with gene expression programs of primary human liver cancer. Specifically, the combinations of imaging features were able to reconstruct 78% of the global gene expression profiles, revealing cell proliferation, liver synthetic function, and patient prognosis [150]. Recently, Miura and coauthors aimed at evaluating the clinicopathological and biological properties of high HCC-correlating ethoxybenzyl-magnetic resonance imaging hyperintensity with gene expression signatures [151]. They identified 53 up-regulated and 71 down-regulated probe sets in the high-HCC group compared with the low-HCC group, showing that clinicopathological and global gene expression analyses revealed low-grade malignancies in high HCCs compared with low HCCs.

3.7. General Considerations

Imaging features listed in Table 1 show that the majority of studies are based on a qualitative and supervised analyses of imaging data, therefore, partially exploiting the full potential of radiomics features as described above. Although the NGS approaches could offer high capabilities for in-depth characterization of tumour biology, most of the radiogenomic studies integrate genomic features by using microarray to assess differential changes in gene expression levels. In fact, the application of this methodology in radiogenomic studies is limited due to the lack of standard references for analysis, it is expensive in terms of cost and time, and Big Data are produced, which require a dedicate storage platform [155]. In addition, Table 1 shows a wide heterogeneity of methods for features generation, as well as the statistical analyses employed, supporting that radiogenomics is still an elusive concept, wherein the standardization of procedures should be addressed.

In this context, an interesting scientific debate begun about the potential utility of radiogenomics compared to separated imaging and genomics approaches in clinical practice [8,9]. One advantage of the integrated approach arises from the limitations of currently available datasets, exploiting its ability to deal with limited and incomplete data to generate meaningful information [8]. Scientific studies regarding the relationship between a few imaging features and a restricted number of gene expressions should not be considered as a pure radiogenomics approach. Radiogenomic studies must combine a large number of quantitative imaging features with a massive genomic signature using computer algorithms. In addition, both radiomics and genomics are needful for the clinical decision making and neither one can replace the other, but their potential can be increased through the interpretation of the two methods to improve the management of cancer patients. Furthermore, the study of mutual relationships between imaging and genomics can provide novel insights for the understanding of neoplastic transformation.

The large number of measured variables can lead to a “fishing expedition”; as discussed in Reference [81], such a poor practice can be avoided with appropriate prior hypotheses about the study. From a statistical point of view, multiple comparisons should be controlled by correcting the accepted criteria for significance, such as the well-known Bonferroni P value correction or the False Discovery Rate (FDR) [118]. In order to translate radiogenomic research to the clinical practice, reproducibility and reliability of the measures and procedures are crucial issues to take into account, requiring the validation on an independent test-retest data set. In addition, a large cohort study gives power to the method validation but, to date, in the published studies an acceptable sample size is lacking. Although genomics has been demonstrated to be a robust analytical tool, the performance of radiomic measurements is still under investigation. Parmaret et al. have investigated the reproducibility of image segmentation and subsequent radiomic feature extraction, showing the better performance of semi-automated techniques [23]. In a recent study Zhao et al. [119] showed that radiomic features are reproducible over a wide range of imaging settings, but they suggest awareness of the importance of properly setting imaging acquisition parameters.

4. Conclusions

As tissue imaging can correlate with tissue pathology, radiomics aims at creating imaging biomarkers that can identify the genomics of a disease, especially cancer, without the need for a biopsy. Integration with genomic data in the last twenty years, including DNA microarrays and RNA-Seq, as allowed the investigation of the relationship between cellular genomics and tissue-scale imaging. The simultaneous use of these apparently different methods to answer a clinical question is called radiogenomics, and is aimed at correlating cancer imaging and gene features.

The increasing amount of imaging and genomic data identified in several types of cancer is encouraging an integrated approach that combines phenotype and genetics for advancing tumour characterization and precision medicine. To this aim, radiogenomics represents a powerful strategy, potentially enabling clinical decision tools to enhance diagnostic accuracy, as well as to measure the response to drug or radiation therapy, leading to an overall improvement of patient management.

However, as with all new diagnostic approaches, it is still in the early phases, limiting its ability to be used to improve cancer patient management requiring standardization of imaging and genomic protocols, image acquisition, and post-processing. The next challenge of radiogenomics will be to identify specific signatures of intra- and inter-tumour heterogeneity in a proper anatomic context. In this way, the radiogenomic analysis of cancer heterogeneity could offer validity by adding and aiding in the assessment of tumour aggressiveness in a variety of clinical settings and oncological outcomes.

Author Contributions: Mariarosaria Incoronato and Marco Aiello conceived, designed, and wrote the manuscript; Teresa Infante, Carlo Cavaliere, Anna Maria Grimaldi, Peppino Mirabelli, and Serena Monti wrote the manuscript; Marco Salvatore critically revised the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Elsheikh, T.M.; Asa, S.L.; Chan, J.K.C.; DeLellis, R.A.; Heffess, C.S.; LiVolsi, V.A.; Wenig, B.M. Interobserver and intraobserver variation among experts in the diagnosis of thyroid follicular lesions with borderline nuclear features of papillary carcinoma. *Am. J. Clin. Pathol.* **2008**, *130*, 736–744. [[CrossRef](#)] [[PubMed](#)]
2. Hariri, A.R.; Weinberger, D.R. Imaging genomics. *Br. Med. Bull.* **2003**, *65*, 259–270. [[CrossRef](#)] [[PubMed](#)]
3. Bai, H.X.; Lee, A.M.; Yang, L.; Zhang, P.; Davatzikos, C.; Maris, J.M.; Diskin, S.J. Imaging genomics in cancer research: Limitations and promises. *Br. J. Radiol.* **2016**, *89*. [[CrossRef](#)] [[PubMed](#)]
4. Obermeyer, Z.; Emanuel, E.J. Predicting the future—Big data, machine learning, and clinical medicine. *N. Engl. J. Med.* **2016**, *375*, 1216–1219. [[CrossRef](#)] [[PubMed](#)]
5. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images are more than pictures, they are data. *Radiology* **2016**, *278*, 563–577. [[CrossRef](#)] [[PubMed](#)]
6. Shen, L.; Kim, S.; Risacher, S.L.; Nho, K.; Swaminathan, S.; West, J.D.; Foroud, T.; Pankratz, N.; Moore, J.H.; Sloan, C.D.; et al. Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: A study of the adni cohort. *Neuroimage* **2010**, *53*, 1051–1063. [[CrossRef](#)] [[PubMed](#)]
7. Aiello, M.; Cavaliere, C.; Salvatore, M. Hybrid PET/MR imaging and brain connectivity. *Front. Neurosci.* **2016**, *10*. [[CrossRef](#)] [[PubMed](#)]
8. Mazurowski, M.A. Radiogenomics: What it is and why it is important. *J. Am. Coll. Radiol.* **2015**, *12*, 862–866. [[CrossRef](#)] [[PubMed](#)]
9. Yang, L.; Bai, H.X.; Lee, A.M. Leveraging imperfect data sets to draw new conclusions: Radiogenomics' true value? *J. Am. Coll. Radiol.* **2016**, *13*, 120–121. [[CrossRef](#)] [[PubMed](#)]
10. Micheel, C.M.; Nass, S.J.; Omenn, G.S. *Evolution of Translational Omics: Lessons Learned and the Path Forward*; National Academies Press: Washington, DC, USA, 2012.
11. Lambin, P.; Rios-Velazquez, E.; Leijenaar, R.; Carvalho, S.; van Stiphout, R.; Granton, P.; Zegers, C.M.L.; Gillies, R.; Boellard, R.; Dekker, A.; et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* **2012**, *48*, 441–446. [[CrossRef](#)] [[PubMed](#)]
12. Xydeas, C.S.; Petrovic, V. Objective image fusion performance measure. *Electron. Lett.* **2000**, *36*, 308–309. [[CrossRef](#)]
13. Kumar, V.; Gu, Y.H.; Basu, S.; Berglund, A.; Eschrich, S.A.; Schabath, M.B.; Forster, K.; Aerts, H.; Dekker, A.; Fenstermacher, D.; et al. Radiomics: The process and the challenges. *Magn. Reson. Imaging* **2012**, *30*, 1234–1248. [[CrossRef](#)] [[PubMed](#)]
14. Anselmo, A.C.; Zhang, M.; Kumar, S.; Vogus, D.R.; Menegatti, S.; Helgeson, M.E.; Mitragotri, S. Elasticity of nanoparticles influences their blood circulation, phagocytosis, endocytosis, and targeting. *ACS Nano* **2015**, *9*, 3169–3177. [[CrossRef](#)] [[PubMed](#)]
15. Galavis, P.E.; Hollensen, C.; Jallow, N.; Paliwal, B.; Jeraj, R. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol.* **2010**, *49*, 1012–1016. [[CrossRef](#)] [[PubMed](#)]
16. Forgacs, A.; Jonsson, H.P.; Dahlbom, M.; Daver, F.; di Franco, M.D.; Opposits, G.; Krizsan, A.K.; Garai, I.; Czernin, J.; Varga, J.; et al. A study on the basic criteria for selecting heterogeneity parameters of F18-FDG PET images. *PLoS ONE* **2016**, *11*, e0164113. [[CrossRef](#)] [[PubMed](#)]

17. Carlier, T.; Bailly, C. State-of-the-art and recent advances in quantification for therapeutic follow-up in oncology using PET. *Front. Med.* **2015**, *2*, 18. [[CrossRef](#)] [[PubMed](#)]
18. Leijenaar, R.T.H.; Carvalho, S.; Velazquez, E.R.; Van Elmpt, W.J.C.; Parmar, C.; Hoekstra, O.S.; Hoekstra, C.J.; Boellaard, R.; Dekker, A.; Gillies, R.J.; et al. Stability of FDG-PET radiomics features: An integrated analysis of test-retest and inter-observer variability. *Acta Oncol.* **2013**, *52*, 1391–1397. [[CrossRef](#)] [[PubMed](#)]
19. Gatenby, R.A.; Grove, O.; Gillies, R.J. Quantitative imaging in cancer evolution and ecology. *Radiology* **2013**, *269*, 8–15. [[CrossRef](#)] [[PubMed](#)]
20. Zhou, M.; Hall, L.; Goldgof, D.; Russo, R.; Balagurunathan, Y.; Gillies, R.; Gatenby, R. Radiologically defined ecological dynamics and clinical outcomes in glioblastoma multiforme: Preliminary results. *Transl. Oncol.* **2014**, *7*, 5–13. [[CrossRef](#)] [[PubMed](#)]
21. Velazquez, E.R.; Aerts, H.; Gu, Y.H.; Goldgof, D.B.; De Ruyscher, D.; Dekker, A.; Korn, R.; Gillies, R.J.; Lambin, P. A semiautomatic CT-based ensemble segmentation of lung tumors: Comparison with oncologists' delineations and with the surgical specimen. *Radiother. Oncol.* **2012**, *105*, 167–173. [[CrossRef](#)] [[PubMed](#)]
22. van Dam, I.E.; de Koste, J.R.V.; Hanna, G.G.; Muirhead, R.; Slotman, B.J.; Senan, S. Improving target delineation on 4-dimensional CT scans in stage I NSCLC using a deformable registration tool. *Radiother. Oncol.* **2010**, *96*, 67–72. [[CrossRef](#)] [[PubMed](#)]
23. Aerts, H. The potential of radiomic-based phenotyping in precision medicine: A review. *JAMA Oncol.* **2016**, *2*, 1636–1642. [[CrossRef](#)] [[PubMed](#)]
24. Nishino, M.; Dahlberg, S.E.; Fulton, L.E.; Digumarthy, S.R.; Hatabu, H.; Johnson, B.E.; Sequist, L.V. Volumetric tumor response and progression in EGFR-mutant NSCLC patients treated with erlotinib or gefitinib. *Acad. Radiol.* **2016**, *23*, 329–336. [[CrossRef](#)] [[PubMed](#)]
25. Nishino, M.; Guo, M.Y.; Jackman, D.M.; DiPiro, P.J.; Yap, J.T.; Ho, T.K.; Hatabu, H.; Janne, P.A.; Van den Abbeele, A.D.; Johnson, B.E. CT tumor volume measurement in advanced non-small-cell lung cancer: Performance characteristics of an emerging clinical tool. *Acad. Radiol.* **2011**, *18*, 54–62. [[CrossRef](#)] [[PubMed](#)]
26. Parmar, C.; Velazquez, E.R.; Leijenaar, R.; Jermoumi, M.; Carvalho, S.; Mak, R.H.; Mitra, S.; Shankar, B.U.; Kikinis, R.; Haibe-Kains, B.; et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS ONE* **2014**, *9*, e102107. [[CrossRef](#)] [[PubMed](#)]
27. O'Connor, J.P.B.; Rose, C.J.; Waterton, J.C.; Carano, R.A.D.; Parker, G.J.M.; Jackson, A. Imaging intratumor heterogeneity: Role in therapy response, resistance, and clinical outcome. *Clin. Cancer Res.* **2015**, *21*, 249–257. [[CrossRef](#)] [[PubMed](#)]
28. American College of Radiology. *Breast Imaging Reporting and Data System Atlas (BI-RADS Atlas)*; American College of Radiology: Reston, VA, USA, 2003; Volume 98.
29. Weinreb, J.C.; Barentsz, J.O.; Choyke, P.L.; Cornud, F.; Haider, M.A.; Macura, K.J.; Margolis, D.; Schnall, M.D.; Shtern, F.; Tempany, C.M.; et al. PI-RADS prostate imaging—Reporting and data system: 2015, version 2. *Eur. Urol.* **2016**, *69*, 16–40. [[CrossRef](#)] [[PubMed](#)]
30. Mazurowski, M.A.; Desjardins, A.; Malof, J.M. Imaging descriptors improve the predictive power of survival models for glioblastoma patients. *Neuro. Oncol.* **2013**, *15*, 1389–1394. [[CrossRef](#)] [[PubMed](#)]
31. Aerts, H.; Grossmann, P.; Tan, Y.Q.; Oxnard, G.G.; Rizvi, N.; Schwartz, L.H.; Zhao, B.S. Defining a radiomic response phenotype: A pilot study using targeted therapy in NSCLC. *Sci. Rep.* **2016**, *6*. [[CrossRef](#)] [[PubMed](#)]
32. Davnall, F.; Yip, C.S.P.; Ljungqvist, G.; Selmi, M.; Ng, F.; Sanghera, B.; Ganeshan, B.; Miles, K.A.; Cook, G.J.; Goh, V. Assessment of tumor heterogeneity: An emerging imaging tool for clinical practice? *Insights Imaging* **2012**, *3*, 573–589. [[CrossRef](#)] [[PubMed](#)]
33. Haralick, R.M.; Shanmugam, K. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *3*, 610–621. [[CrossRef](#)]
34. Galloway, M.M. Texture analysis using gray level run lengths. *Comput. Graph. Image Process.* **1975**, *4*, 172–179. [[CrossRef](#)]
35. Amadasun, M.; King, R. Textural features corresponding to textural properties. *IEEE Trans. Syst. Man Cybern.* **1989**, *19*, 1264–1274. [[CrossRef](#)]
36. Aerts, H.; Velazquez, E.R.; Leijenaar, R.T.H.; Parmar, C.; Grossmann, P.; Carvalho, S.; Bussink, J.; Monshouwer, R.; Haibe-Kains, B.; Rietveld, D.; et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* **2014**, *5*. [[CrossRef](#)]
37. Fogel, I.; Sagi, D. Gabor filters as texture discriminator. *Biol. Cybern.* **1989**, *61*, 103–113. [[CrossRef](#)]

38. Poustka, A.; Pohl, T.; Barlow, D.; Zehetner, G.; Craig, A.; Michiels, F.; Ehrich, E.; Frischauf, A.-M.; Lehrach, H. Molecular approaches to mammalian genetics. *Cold Spring Harb. Symp. Quant. Biol.* **1986**, *51*, 131–139. [[CrossRef](#)] [[PubMed](#)]
39. Cantor, C.R.; Mirzabekov, A.; Southern, E. Report on the sequencing by hybridization workshop. *Genomics* **1992**, *13*, 1378–1383. [[CrossRef](#)]
40. Marzancola, M.G.; Sedighi, A.; Li, P.C.H. DNA microarray-based diagnostics. *Methods Mol. Biol.* **2016**, *1368*, 161–178. [[PubMed](#)]
41. Jaksik, R.; Iwanaszko, M.; Rzeszowska-Wolny, J.; Kimmel, M. Microarray experiments and factors which affect their reliability. *Biol. Direct* **2015**, *10*. [[CrossRef](#)] [[PubMed](#)]
42. Schena, M.; Shalon, D.; Davis, R.W.; Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary dna microarray. *Science* **1995**, *270*, 467–470. [[CrossRef](#)] [[PubMed](#)]
43. Jones, P.A.; Baylin, S.B. The epigenomics of cancer. *Cell* **2007**, *128*, 683–692. [[CrossRef](#)] [[PubMed](#)]
44. Zeller, C.; Masrour, N.; Patel, N.; Dai, W.; Wilhelm-Benartzi, C.; Brown, R. DNA methylation profiling using infinium methylation assay. *Bio Protocol*. **2013**, *3*. [[CrossRef](#)]
45. Meldrum, C.; Doyle, M.A.; Tohill, R.W. Next-generation sequencing for cancer diagnostics: A practical perspective. *Clin. Biochem. Rev.* **2011**, *32*, 177–195. [[PubMed](#)]
46. Bahassi, E.; Stambrook, P.J. Next-generation sequencing technologies: Breaking the sound barrier of human genetics. *Mutagenesis* **2014**, *29*, 303–310. [[CrossRef](#)] [[PubMed](#)]
47. Metzker, M.L. Applications of next-generation sequencing technologies—The next generation. *Nat. Rev. Genet.* **2010**, *11*, 31–46. [[CrossRef](#)] [[PubMed](#)]
48. Voelkerding, K.V.; Dames, S.A.; Durtschi, J.D. Next-generation sequencing: From basic research to diagnostics. *Clin. Chem.* **2009**, *55*, 641–658. [[CrossRef](#)] [[PubMed](#)]
49. Wade, J.T. Mapping transcription regulatory networks with ChIP-Seq and RNA-Seq. *Adv. Exp. Med. Biol.* **2015**, *883*, 119–134. [[PubMed](#)]
50. Sun, Z.F.; Cunningham, J.; Slager, S.; Kocher, J.P. Base resolution methylome profiling: Considerations in platform selection, data preprocessing and analysis. *Epigenomics* **2015**, *7*, 813–828. [[CrossRef](#)] [[PubMed](#)]
51. Costa, V.; Angelini, C.; De Feis, I.; Ciccodicola, A. Uncovering the complexity of transcriptomes with RNA-Seq. *J. Biomed. Biotechnol.* **2010**, *2010*. [[CrossRef](#)] [[PubMed](#)]
52. Sager, M.; Yeat, N.C.; Pajaro-Van der Stadt, S.; Lin, C.; Ren, Q.Y.; Lin, J. Transcriptomics in cancer diagnostics: Developments in technology, clinical research and commercialization. *Expert Rev. Mol. Diagn.* **2015**, *15*, 1589–1603. [[CrossRef](#)] [[PubMed](#)]
53. Pan, Q.; Shai, O.; Lee, L.J.; Frey, J.; Blencowe, B.J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **2008**, *40*, 1413–1415. [[CrossRef](#)] [[PubMed](#)]
54. Guttman, M.; Garber, M.; Levin, J.Z.; Donaghey, J.; Robinson, J.; Adiconis, X.; Fan, L.; Koziol, M.J.; Gnirke, A.; Nusbaum, C.; et al. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat. Biotechnol.* **2010**, *28*, 503–510. [[CrossRef](#)] [[PubMed](#)]
55. Rienzo, M.; Costa, V.; Scarpato, M.; Schiano, C.; Casamassimi, A.; Grimaldi, V.; Ciccodicola, A.; Napoli, C. RNA-Seq for the identification of novel mediator transcripts in endothelial progenitor cells. *Gene* **2014**, *547*, 98–105. [[CrossRef](#)] [[PubMed](#)]
56. Degner, J.F.; Marioni, J.C.; Pai, A.A.; Pickrell, J.K.; Nkadori, E.; Gilad, Y.; Pritchard, J.K. Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. *Bioinformatics* **2009**, *25*, 3207–3212. [[CrossRef](#)] [[PubMed](#)]
57. Edgren, H.; Murumagi, A.; Kangaspeska, S.; Nicorici, D.; Hongisto, V.; Kleivi, K.; Rye, I.H.; Nyberg, S.; Wolf, M.; Borresen-Dale, A.L.; et al. Identification of fusion genes in breast cancer by paired-end RNA-sequencing. *Genome Biol.* **2011**, *12*. [[CrossRef](#)] [[PubMed](#)]
58. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nat. Methods* **2008**, *5*, 621–628. [[CrossRef](#)] [[PubMed](#)]
59. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **2009**, *10*, 57–63. [[CrossRef](#)] [[PubMed](#)]
60. Peng, Z.Y.; Cheng, Y.B.; Tan, B.C.M.; Kang, L.; Tian, Z.J.; Zhu, Y.K.; Zhang, W.W.; Liang, Y.; Hu, X.D.; Tan, X.M.; et al. Comprehensive analysis of RNA-seq data reveals extensive RNA editing in a human transcriptome. *Nat. Biotechnol.* **2012**, *30*, 253–260. [[CrossRef](#)] [[PubMed](#)]

61. Cloonan, N.; Forrest, A.R.R.; Kolle, G.; Gardiner, B.B.A.; Faulkner, G.J.; Brown, M.K.; Taylor, D.F.; Steptoe, A.L.; Wani, S.; Bethel, G.; et al. Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat. Methods* **2008**, *5*, 613–619. [[CrossRef](#)] [[PubMed](#)]
62. Marioni, J.C.; Mason, C.E.; Mane, S.M.; Stephens, M.; Gilad, Y. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* **2008**, *18*, 1509–1517. [[CrossRef](#)] [[PubMed](#)]
63. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
64. Wang, L.G.; Wang, S.Q.; Li, W. Rseqc: Quality control of RNA-seq experiments. *Bioinformatics* **2012**, *28*, 2184–2185. [[CrossRef](#)] [[PubMed](#)]
65. Garcia-Alcalde, F.; Okonechnikov, K.; Carbonell, J.; Cruz, L.M.; Gotz, S.; Tarazona, S.; Dopazo, J.; Meyer, T.F.; Conesa, A. Qualimap: Evaluating next-generation sequencing alignment data. *Bioinformatics* **2012**, *28*, 2678–2679. [[CrossRef](#)] [[PubMed](#)]
66. Xie, Y.L.; Wu, G.X.; Tang, J.B.; Luo, R.B.; Patterson, J.; Liu, S.L.; Huang, W.H.; He, G.Z.; Gu, S.C.; Li, S.K.; et al. Soapdenovo-trans: De novo transcriptome assembly with short RNA-Seq reads. *Bioinformatics* **2014**, *30*, 1660–1666. [[CrossRef](#)] [[PubMed](#)]
67. Schulz, M.H.; Zerbino, D.R.; Vingron, M.; Birney, E. Oases: Robust de novo RNA-Seq assembly across the dynamic range of expression levels. *Bioinformatics* **2012**, *28*, 1086–1092. [[CrossRef](#)] [[PubMed](#)]
68. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.D.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)] [[PubMed](#)]
69. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M.; et al. De novo transcript sequence reconstruction from RNA-Seq using the trinity platform for reference generation and analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512. [[CrossRef](#)] [[PubMed](#)]
70. Steijger, T.; Abril, J.F.; Engstrom, P.G.; Kokocinski, F.; Hubbard, T.J.; Guigo, R.; Harrow, J.; Bertone, P.; Consortium, R. Assessment of transcript reconstruction methods for RNA-Seq. *Nat. Methods* **2013**, *10*, 1177–1184. [[CrossRef](#)] [[PubMed](#)]
71. Boley, N.; Stoiber, M.H.; Booth, B.W.; Wan, K.H.; Hoskins, R.A.; Bickel, P.J.; Celniker, S.E.; Brown, J.B. Genome-guided transcript assembly by integrative analysis of RNA sequence data. *Nat. Biotechnol.* **2014**, *32*, 341–346. [[CrossRef](#)] [[PubMed](#)]
72. Roberts, A.; Pimentel, H.; Trapnell, C.; Pachter, L. Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics* **2011**, *27*, 2325–2329. [[CrossRef](#)] [[PubMed](#)]
73. Mezlini, A.M.; Smith, E.J.M.; Fiume, M.; Buske, O.; Savich, G.L.; Shah, S.; Aparicio, S.; Chiang, D.Y.; Goldenberg, A.; Brudno, M. Ireckon: Simultaneous isoform discovery and abundance estimation from RNA-Seq data. *Genome Res.* **2013**, *23*, 519–529. [[CrossRef](#)] [[PubMed](#)]
74. Pertea, M.; Pertea, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. Stringtie enables improved reconstruction of a transcriptome from RNA-Seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295. [[CrossRef](#)] [[PubMed](#)]
75. Engstrom, P.G.; Steijger, T.; Sipos, B.; Grant, G.R.; Kahles, A.; Ratsch, G.; Goldman, N.; Hubbard, T.J.; Harrow, J.; Guigo, R.; et al. Systematic evaluation of spliced alignment programs for RNA-Seq data. *Nat. Methods* **2013**, *10*, 1185–1191. [[CrossRef](#)] [[PubMed](#)]
76. Stanke, M.; Keller, O.; Gunduz, I.; Hayes, A.; Waack, S.; Morgenstern, B. Augustus: Ab initio prediction of alternative transcripts. *Nucleic Acids Res.* **2006**, *34*, 435–439. [[CrossRef](#)] [[PubMed](#)]
77. Hiller, D.; Wong, W.H. Simultaneous isoform discovery and quantification from RNA-Seq. *Stat. Biosci.* **2013**, *5*, 100–118. [[CrossRef](#)] [[PubMed](#)]
78. Anders, S.; Pyl, P.T.; Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**, *31*, 166–169. [[CrossRef](#)] [[PubMed](#)]
79. Liao, Y.; Smyth, G.K.; Shi, W. Featurecounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014**, *30*, 923–930. [[CrossRef](#)] [[PubMed](#)]
80. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)] [[PubMed](#)]
81. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. Star: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21. [[CrossRef](#)] [[PubMed](#)]

82. Li, H.; Durbin, R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]
83. Wu, A.R.; Neff, N.F.; Kalisky, T.; Dalerba, P.; Treutlein, B.; Rothenberg, M.E.; Mburu, F.M.; Mantalas, G.L.; Sim, S.; Clarke, M.F.; et al. Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* **2014**, *11*, 41–46. [[CrossRef](#)] [[PubMed](#)]
84. Conesa, A.; Madrigal, P.; Tarazona, S.; Gomez-Cabrero, D.; Cervera, A.; McPherson, A.; Szczesniak, M.W.; Gaffney, D.J.; Elo, L.L.; Zhang, X.G.; et al. A survey of best practices for RNA-Seq data analysis. *Genome Biol.* **2016**, *17*. [[CrossRef](#)] [[PubMed](#)]
85. Wilkerson, M.D.; Cabanski, C.R.; Sun, W.; Hoadley, K.A.; Walter, V.; Mose, L.E.; Troester, M.A.; Hammerman, P.S.; Parker, J.S.; Perou, C.M.; et al. Integrated RNA and DNA sequencing improves mutation detection in low purity tumors. *Nucleic Acids Res.* **2014**, *42*. [[CrossRef](#)] [[PubMed](#)]
86. Ke, R.Q.; Mignardi, M.; Pacureanu, A.; Svedlund, J.; Botling, J.; Wahlby, C.; Nilsson, M. In situ sequencing for rna analysis in preserved tissue and cells. *Nat. Methods* **2013**, *10*, 857–860. [[CrossRef](#)] [[PubMed](#)]
87. Hrdlickova, R.; Toloue, M.; Tian, B. RNA-Seq methods for transcriptome analysis. *Wiley Interdiscip. Rev. RNA* **2017**, *8*. [[CrossRef](#)] [[PubMed](#)]
88. Ramos-Vara, J.A. Principles and methods of immunohistochemistry. *Drug Saf. Eval. Methods Protoc.* **2011**, *691*, 83–96.
89. Kuo, M.D.; Jamshidi, N. Behind the numbers: Decoding molecular phenotypes with radiogenomics-guiding principles and technical considerations. *Radiology* **2014**, *270*, 320–325. [[CrossRef](#)] [[PubMed](#)]
90. Zhang, B.; Chang, K.; Ramkissoon, S.; Tanguturi, S.; Bi, W.L.; Reardon, D.A.; Ligon, K.L.; Alexander, B.M.; Wen, P.Y.; Huang, R.Y. Multimodal MRI features predict isocitrate dehydrogenase genotype in high-grade gliomas. *Neuro. Oncol.* **2017**, *19*, 109–117. [[CrossRef](#)] [[PubMed](#)]
91. Zhu, Y.T.; Li, H.; Guo, W.T.; Drukker, K.; Lan, L.; Giger, M.L.; Ji, Y. Deciphering genomic underpinnings of quantitative MRI-based radiomic phenotypes of invasive breast carcinoma. *Sci. Rep.* **2015**, *5*. [[CrossRef](#)] [[PubMed](#)]
92. Friedman, J.H. On bias, variance, 0/1—Loss, and the curse-of-dimensionality. *Data Min. Know. Discov.* **1997**, *1*, 55–77. [[CrossRef](#)]
93. Balagurunathan, Y.; Kumar, V.; Gu, Y.H.; Kim, J.; Wang, H.; Liu, Y.; Goldgof, D.B.; Hall, L.O.; Korn, R.; Zhao, B.S.; et al. Test-retest reproducibility analysis of lung CT image features. *J. Digit. Imaging* **2014**, *27*, 805–823. [[CrossRef](#)] [[PubMed](#)]
94. Kuo, M.D.; Gollub, J.; Sirlin, C.B.; Ooi, C.; Chen, X. Radiogenomic analysis to identify imaging phenotypes associated with drug response gene expression programs in hepatocellular carcinoma. *J. Vasc. Interv. Radiol.* **2007**, *18*, 821–831. [[CrossRef](#)] [[PubMed](#)]
95. Wu, W.M.; Parmar, C.; Grossmann, P.; Quackenbush, J.; Lambin, P.; Bussink, J.; Mak, R.; Aerts, H. Exploratory study to identify radiomics classifiers for lung cancer histology. *Front. Oncol.* **2016**, *6*. [[CrossRef](#)] [[PubMed](#)]
96. Parmar, C.; Grossmann, P.; Bussink, J.; Lambin, P.; Aerts, H. Machine learning methods for quantitative radiomic biomarkers. *Sci. Rep.* **2015**, *5*. [[CrossRef](#)] [[PubMed](#)]
97. Guo, W.T.; Li, H.; Zhu, Y.T.; Lan, L.; Yang, S.J.; Drukker, K.; Morris, E.; Burnside, E.; Whitman, G.; Giger, M.L.; et al. Prediction of clinical phenotypes in invasive breast carcinomas from the integration of radiomics and genomics data. *J. Med. Imaging* **2015**, *2*. [[CrossRef](#)] [[PubMed](#)]
98. Wu, T.T.; Chen, Y.F.; Hastie, T.; Sobel, E.; Lange, K. Genome-wide association analysis by lasso penalized logistic regression. *Bioinformatics* **2009**, *25*, 714–721. [[CrossRef](#)] [[PubMed](#)]
99. Jamshidi, N.; Jonasch, E.; Zapala, M.; Korn, R.L.; Aganovic, L.; Zhao, H.J.; Sitaram, R.T.; Tibshirani, R.J.; Banerjee, S.; Brooks, J.D.; et al. The radiogenomic risk score: Construction of a prognostic quantitative, noninvasive image-based molecular assay for renal cell carcinoma. *Radiology* **2015**, *277*, 114–123. [[CrossRef](#)] [[PubMed](#)]
100. Leung, M.K.K.; Delong, A.; Alipanahi, B.; Frey, B.J. Machine learning in genomic medicine: A review of computational problems and data sets. *Proc. IEEE* **2016**, *104*, 176–197. [[CrossRef](#)]
101. Deng, L.; Yu, D. Deep learning: Methods and applications. *Found. Trends Signal Proc.* **2014**, *7*, 197–387. [[CrossRef](#)]

102. Panth, K.M.; Leijenaar, R.T.H.; Carvalho, S.; Lieuwes, N.G.; Yaromina, A.; Dubois, L.; Lambin, P. Is there a causal relationship between genetic changes and radiomics-based image features? An in vivo preclinical experiment with doxycycline inducible GADD34 tumor cells. *Radiother. Oncol.* **2015**, *116*, 462–466. [[CrossRef](#)] [[PubMed](#)]
103. Grimm, L.J.; Zhang, J.; Mazurowski, M.A. Computational approach to radiogenomics of breast cancer: Luminal A and luminal B molecular subtypes are associated with imaging features on routine breast mri extracted using computer vision algorithms. *J. Magn. Reson. Imaging* **2015**, *42*, 902–907. [[CrossRef](#)] [[PubMed](#)]
104. Mazurowski, M.A.; Zhang, J.; Grimm, L.J.; Yoon, S.C.; Silber, J.I. Radiogenomic analysis of breast cancer: Luminal b molecular subtype is associated with enhancement dynamics at MR imaging. *Radiology* **2014**, *273*, 365–372. [[CrossRef](#)] [[PubMed](#)]
105. Sung, J.S.; Jochelson, M.S.; Brennan, S.; Joo, S.; Wen, Y.H.; Moskowitz, C.; Zheng, J.T.; Dershaw, D.D.; Morris, E.A. MR imaging features of triple-negative breast cancers. *Breast J.* **2013**, *19*, 643–649. [[CrossRef](#)] [[PubMed](#)]
106. Sutton, E.J.; Oh, J.H.; Dashevsky, B.Z.; Veeraraghavan, H.; Apte, A.P.; Thakur, S.B.; Deasy, J.O.; Morris, E.A. Breast cancer subtype intertumor heterogeneity: MRI-based features predict results of a genomic assay. *J. Magn. Reson. Imaging* **2015**, *42*, 1398–1406. [[CrossRef](#)] [[PubMed](#)]
107. Sutton, E.J.; Dashevsky, B.Z.; Oh, J.H.; Veeraraghavan, H.; Apte, A.P.; Thakur, S.B.; Morris, E.A.; Deasy, J.O. Breast cancer molecular subtype classifier that incorporates MRI features. *J. Magn. Reson. Imaging* **2016**, *44*, 122–129. [[CrossRef](#)] [[PubMed](#)]
108. Uematsu, T.; Kasami, M.; Yuen, S. Triple-negative breast cancer: Correlation between MR imaging and pathologic findings. *Radiology* **2009**, *250*, 638–647. [[CrossRef](#)] [[PubMed](#)]
109. Yamamoto, S.; Maki, D.D.; Korn, R.L.; Kuo, M.D. Radiogenomic analysis of breast cancer using MRI: A preliminary study to define the landscape. *Am. J. Roentgenol.* **2012**, *199*, 654–663. [[CrossRef](#)] [[PubMed](#)]
110. Yamamoto, S.; Han, W.; Kim, Y.; Du, L.T.; Jamshidi, N.; Huang, D.S.; Kim, J.H.; Kuo, M.D. Breast cancer: Radiogenomic biomarker reveals associations among dynamic contrast-enhanced MR imaging, long noncoding rna, and metastasis. *Radiology* **2015**, *275*, 384–392. [[CrossRef](#)] [[PubMed](#)]
111. Cipolla, V.; Santucci, D.; Guerrieri, D.; Drudi, F.M.; Meggiorini, M.L.; de Felice, C. Correlation between 3T apparent diffusion coefficient values and grading of invasive breast carcinoma. *Eur. J. Radiol.* **2014**, *83*, 2144–2150. [[CrossRef](#)] [[PubMed](#)]
112. Molinari, C.; Clauser, P.; Girometti, R.; Linda, A.; Cimino, E.; Puglisi, F.; Zuiani, C.; Bazzocchi, M. MR mammography using diffusion-weighted imaging in evaluating breast cancer: A correlation with proliferation index. *Radiol. Med.* **2015**, *120*, 911–918. [[CrossRef](#)] [[PubMed](#)]
113. An, Y.Y.; Kim, S.H.; Kang, B.J.; Park, C.S.; Jung, N.Y.; Kim, J.Y. Breast cancer in very young women (<30 years): Correlation of imaging features with clinicopathological features and immunohistochemical subtypes. *Eur. J. Radiol.* **2015**, *84*, 1894–1902. [[CrossRef](#)] [[PubMed](#)]
114. Bae, M.S.; Seo, M.; Kim, K.G.; Park, I.A.; Moon, W.K. Quantitative mri morphology of invasive breast cancer: Correlation with immunohistochemical biomarkers and subtypes. *Acta Radiol.* **2015**, *56*, 269–275. [[CrossRef](#)] [[PubMed](#)]
115. Shin, H.J.; Park, J.Y.; Shin, K.C.; Kim, H.H.; Cha, J.H.; Chae, E.Y.; Choi, W.J. Characterization of tumor and adjacent peritumoral stroma in patients with breast cancer using high-resolution diffusion-weighted imaging: Correlation with pathologic biomarkers. *Eur. J. Radiol.* **2016**, *85*, 1004–1011. [[CrossRef](#)] [[PubMed](#)]
116. Kitajim, K.; Yamano, T.; Fukushima, K.; Miyoshi, Y.; Hirota, S.; Kawanaka, Y.; Miya, M.; Doi, H.; Yamakado, K. Correlation of the SUVmax of FDG-PET and ADC values of diffusion-weighted mr imaging with pathologic prognostic factors in breast carcinoma. *Eur. J. Radiol.* **2016**, *85*, 943–949. [[CrossRef](#)] [[PubMed](#)]
117. Koo, H.R.; Park, J.S.; Kang, K.W.; Han, W.; Park, I.A.; Moon, W.K. Correlation between F-18-FDG uptake on PET/CT and prognostic factors in triple-negative breast cancer. *Eur. Radiol.* **2015**, *25*, 3314–3321. [[CrossRef](#)] [[PubMed](#)]
118. Koo, H.R.; Park, J.S.; Kang, K.W.; Cho, N.; Chang, J.M.; Bae, M.S.; Kim, W.H.; Lee, S.H.; Kim, M.Y.; Kim, J.Y.; et al. F-18-FDG uptake in breast cancer correlates with immunohistochemically defined subtypes. *Eur. Radiol.* **2014**, *24*, 610–618. [[CrossRef](#)] [[PubMed](#)]
119. Tchou, J.; Sonnad, S.S.; Bergey, M.R.; Basu, S.; Tomaszewski, J.; Alavi, A.; Schnall, M. Degree of tumor FDG uptake correlates with proliferation index in triple negative breast cancer. *Mol. Imaging Biol.* **2010**, *12*, 657–662. [[CrossRef](#)] [[PubMed](#)]

120. Yamaguchi, K.; Abe, H.; Newstead, G.M.; Egashira, R.; Nakazono, T.; Imaizumi, T.; Irie, H. Intratumoral heterogeneity of the distribution of kinetic parameters in breast cancer: Comparison based on the molecular subtypes of invasive breast cancer. *Breast Cancer* **2015**, *22*, 496–502. [[CrossRef](#)] [[PubMed](#)]
121. Mehta, S.; Hughes, N.P.; Li, S.; Jubb, A.; Adams, R.; Lord, S.; Koumakis, L.; van Stiphout, R.; Padhani, A.; Makris, A.; et al. Radiogenomics monitoring in breast cancer identifies metabolism and immune checkpoints as early actionable mechanisms of resistance to anti-angiogenic treatment. *Ebiomedicine* **2016**, *10*, 109–116. [[CrossRef](#)] [[PubMed](#)]
122. Osborne, J.R.; Port, E.; Gonen, M.; Doane, A.S.; Yeung, H.; Gerald, W.; Cook, J.B.; Larson, S. F-18-FDG PET of locally invasive breast cancer and association of estrogen receptor status with standardized uptake value: Microarray and immunohistochemical analysis. *J. Nucl. Med.* **2010**, *51*, 543–550. [[CrossRef](#)] [[PubMed](#)]
123. Palaskas, N.; Larson, S.M.; Schultz, N.; Komisopoulou, E.; Wong, J.; Rohle, D.; Campos, C.; Yannuzzi, N.; Osborne, J.R.; Linkov, I.; et al. F-18-fluorodeoxy-glucose positron emission tomography marks MYC-overexpressing human basal-like breast cancers. *Cancer Res.* **2011**, *71*, 5164–5174. [[CrossRef](#)] [[PubMed](#)]
124. Diehn, M.; Nardini, C.; Wang, D.S.; McGovern, S.; Jayaraman, M.; Liang, Y.; Alclape, K.; Cha, S.; Kuo, M.D. Identification of noninvasive imaging surrogates for brain tumor gene-expression modules. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5213–5218. [[CrossRef](#)] [[PubMed](#)]
125. Zinn, P.O.; Majadan, B.; Sathyan, P.; Singh, S.K.; Majumder, S.; Jolesz, F.A.; Colen, R.R. Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme. *PLoS ONE* **2011**, *6*, e25451. [[CrossRef](#)] [[PubMed](#)]
126. Jamshidi, N.; Diehn, M.; Bredel, M.; Kuo, M.D. Illuminating radiogenomic characteristics of glioblastoma multiforme through integration of MR imaging, messenger RNA expression, and DNA copy number variation. *Radiology* **2014**, *270*, 212–222. [[CrossRef](#)] [[PubMed](#)]
127. Gutman, D.A.; Dunn, W.D.; Grossmann, P.; Cooper, L.A.D.; Holder, C.A.; Ligon, K.L.; Alexander, B.M.; Aerts, H. Somatic mutations associated with MRI-derived volumetric features in glioblastoma. *Neuroradiology* **2015**, *57*, 1227–1237. [[CrossRef](#)] [[PubMed](#)]
128. Rao, A.; Rao, G.; Gutman, D.A.; Flanders, A.E.; Hwang, S.N.; Rubin, D.L.; Colen, R.R.; Zinn, P.O.; Jain, R.; Wintermark, M.; et al. A combinatorial radiographic phenotype may stratify patient survival and be associated with invasion and proliferation characteristics in glioblastoma. *J. Neurosurg.* **2016**, *124*, 1008–1017. [[CrossRef](#)] [[PubMed](#)]
129. Hu, L.S.; Ning, S.; Eschbacher, J.M.; Baxter, L.C.; Gaw, N.; Ranjbar, S.; Plasencia, J.; Dueck, A.C.; Peng, S.; Smith, K.A. Radiogenomics to characterize regional genetic heterogeneity in glioblastoma. *Neuro. Oncol.* **2017**, *19*, 128–137. [[CrossRef](#)] [[PubMed](#)]
130. Gevaert, O.; Mitchell, L.A.; Achrol, A.S.; Xu, J.J.; Echegaray, S.; Steinberg, G.K.; Cheshier, S.H.; Napel, S.; Zaharchuk, G.; Plevritis, S.K. Glioblastoma multiforme: Exploratory radiogenomic analysis by using quantitative image features. *Radiology* **2014**, *273*, 168–174. [[CrossRef](#)] [[PubMed](#)]
131. Heiland, D.H.; Demerath, T.; Kellner, E.; Kiselev, V.G.; Pfeifer, D.; Schnell, O.; Staszewski, O.; Urbach, H.; Weyerbrock, A.; Mader, I. Molecular differences between cerebral blood volume and vessel size in glioblastoma multiforme. *Oncotarget* **2016**. [[CrossRef](#)] [[PubMed](#)]
132. Kickingereder, P.; Bonekamp, D.; Nowosielski, M.; Kratz, A.; Sill, M.; Burth, S.; Wick, A.; Eidel, O.; Schlemmer, H.P.; Radbruch, A.; et al. Radiogenomics of glioblastoma: Machine learning-based classification of molecular characteristics by using multiparametric and multiregional MR imaging features. *Radiology* **2016**, *281*, 907–918. [[CrossRef](#)] [[PubMed](#)]
133. Barajas, R.F.; Hodgson, J.G.; Chang, J.S.; Vandenberg, S.R.; Yeh, R.F.; Parsa, A.T.; McDermott, M.W.; Berger, M.S.; Dillon, W.P.; Cha, S. Glioblastoma multiforme regional genetic and cellular expression patterns: Influence on anatomic and physiologic MR imaging. *Radiology* **2010**, *254*, 564–576. [[CrossRef](#)] [[PubMed](#)]
134. Pope, W.B.; Chen, J.H.; Dong, J.; Carlson, M.R.J.; Perlina, A.; Cloughesy, T.F.; Liau, L.M.; Mischel, P.S.; Nghiemphu, P.; Lai, A.; et al. Relationship between gene expression and enhancement in glioblastoma multiforme: Exploratory DNA microarray analysis. *Radiology* **2008**, *249*, 268–277. [[CrossRef](#)] [[PubMed](#)]
135. Zinn, P.O.; Sathyan, P.; Mahajan, B.; Bruyere, J.; Hegi, M.; Majumder, S.; Colen, R.R. A novel volume-age-KPS (VAK) glioblastoma classification identifies a prognostic cognate microRNA-gene signature. *PLoS ONE* **2012**, *7*, e41522. [[CrossRef](#)] [[PubMed](#)]

136. Carlson, M.R.J.; Pope, W.B.; Horvath, S.; Braunstein, J.G.; Nghiemphu, P.; Tso, C.L.; Mellinghoff, I.; Lai, A.; Liao, L.M.; Mischel, P.S.; et al. Relationship between survival and edema in malignant gliomas: Role of vascular endothelial growth factor and neuronal pentraxin 2. *Clin. Cancer Res.* **2007**, *13*, 2592–2598. [[CrossRef](#)] [[PubMed](#)]
137. Colen, R.R.; Vangel, M.; Wang, J.X.; Gutman, D.A.; Hwang, S.N.; Wintermark, M.; Jain, R.; Jilwan-Nicolas, M.; Chen, J.Y.; Raghavan, P.; et al. Imaging genomic mapping of an invasive mri phenotype predicts patient outcome and metabolic dysfunction: A TCGA glioma phenotype research group project. *BMC Med. Genom.* **2014**, *7*. [[CrossRef](#)] [[PubMed](#)]
138. Jain, R.; Poisson, L.; Narang, J.; Scarpace, L.; Rosenblum, M.L.; Rempel, S.; Mikkelsen, T. Correlation of perfusion parameters with genes related to angiogenesis regulation in glioblastoma: A feasibility study. *Am. J. Neuroradiol.* **2012**, *33*, 1343–1348. [[CrossRef](#)] [[PubMed](#)]
139. Naeni, K.M.; Pope, W.B.; Cloughesy, T.F.; Harris, R.J.; Lai, A.; Eskin, A.; Chowdhury, R.; Phillips, H.S.; Nghiemphu, P.L.; Behbahanian, Y.; et al. Identifying the mesenchymal molecular subtype of glioblastoma using quantitative volumetric analysis of anatomic magnetic resonance images. *Neuro Oncol.* **2013**, *15*, 626–634. [[CrossRef](#)] [[PubMed](#)]
140. Nicolasjilwan, M.; Hu, Y.; Yan, C.H.; Meerzaman, D.; Holder, C.A.; Gutman, D.; Jain, R.; Colen, R.; Rubin, D.L.; Zinn, P.O.; et al. Addition of MR imaging features and genetic biomarkers strengthens glioblastoma survival prediction in TCGA patients. *J. Neuroradiol.* **2015**, *42*, 212–221. [[CrossRef](#)] [[PubMed](#)]
141. Pope, W.B.; Prins, R.M.; Thomas, M.A.; Nagarajan, R.; Yen, K.E.; Bittinger, M.A.; Salamon, N.; Chou, A.P.; Yong, W.H.; Soto, H.; et al. Non-invasive detection of 2-hydroxyglutarate and other metabolites in IDH1 mutant glioma patients using magnetic resonance spectroscopy. *J. Neuro Oncol.* **2012**, *107*, 197–205. [[CrossRef](#)] [[PubMed](#)]
142. Demerath, T.; Simon-Gabriel, C.P.; Kellner, E.; Schwarzwald, R.; Lange, T.; Heiland, D.H.; Reinacher, P.; Staszewski, O.; Mast, H.; Kiselev, V.G.; et al. Mesoscopic imaging of glioblastomas: Are diffusion, perfusion and spectroscopic measures influenced by the radiogenetic phenotype? *Neuroradiol. J.* **2017**, *30*, 36–47. [[CrossRef](#)] [[PubMed](#)]
143. Cui, Y.; Ren, S.; Tha, K.K.; Wu, J.; Shirato, H.; Li, R. Volume of high-risk intratumoral subregions at multi-parametric MR imaging predicts overall survival and complements molecular analysis of glioblastoma. *Eur. Radiol.* **2017**. [[CrossRef](#)] [[PubMed](#)]
144. Gevaert, O.; Echegaray, S.; Khuong, A.; Hoang, C.D.; Shrager, J.B.; Jensen, K.C.; Berry, G.J.; Guo, H.H.; Lau, C.; Plevritis, S.K.; et al. Predictive radiogenomics modeling of EGFR mutation status in lung cancer. *Sci. Rep.* **2017**, *7*. [[CrossRef](#)] [[PubMed](#)]
145. Gevaert, O.; Xu, J.J.; Hoang, C.D.; Leung, A.N.; Xu, Y.; Quon, A.; Rubin, D.L.; Napel, S.; Plevritis, S.K. Non-small cell lung cancer: Identifying prognostic imaging biomarkers by leveraging public gene expression microarray data-methods and preliminary results. *Radiology* **2012**, *264*, 387–396. [[CrossRef](#)] [[PubMed](#)]
146. Nair, V.S.; Gevaert, O.; Davidzon, G.; Plevritis, S.K.; West, R. NF- κ B protein expression associates with F-18-FDG PET tumor uptake in non-small cell lung cancer: A radiogenomics validation study to understand tumor metabolism. *Lung Cancer* **2014**, *83*, 189–196. [[CrossRef](#)] [[PubMed](#)]
147. Karlo, C.A.; Di Paolo, P.L.; Chaim, J.; Hakimi, A.A.; Ostrovskaya, I.; Russo, P.; Hricak, H.; Motzer, R.; Hsieh, J.J.; Akin, O. Radiogenomics of clear cell renal cell carcinoma: Associations between CT imaging features and mutations. *Radiology* **2014**, *270*, 464–471. [[CrossRef](#)] [[PubMed](#)]
148. Shinagare, A.B.; Vikram, R.; Jaffe, C.; Akin, O.; Kirby, J.; Huang, E.; Freymann, J.; Sainani, N.I.; Sadow, C.A.; Bathala, T.K.; et al. Radiogenomics of clear cell renal cell carcinoma: Preliminary findings of the cancer genome Atlas-Renal Cell Carcinoma (TCGA-RCC) imaging research group. *Abdom. Imaging* **2015**, *40*, 1684–1692. [[CrossRef](#)] [[PubMed](#)]
149. Chen, X.; Zhou, Z.; Thomas, K.; Wang, J. Mo-de-207b-05: Predicting gene mutations in renal cell carcinoma based on CT imaging features: Validation using TCGA-TCIA datasets. *Med. Phys.* **2016**, *43*, 3705. [[CrossRef](#)]
150. Segal, E.; Sirlin, C.B.; Ooi, C.; Adler, A.S.; Gollub, J.; Chen, X.; Chan, B.K.; Matcuk, G.R.; Barry, C.T.; Chang, H.Y.; et al. Decoding global gene expression programs in liver cancer by noninvasive imaging. *Nat. Biotechnol.* **2007**, *25*, 675–680. [[CrossRef](#)] [[PubMed](#)]

151. Miura, T.; Ban, D.; Tanaka, S.; Mogushi, K.; Kudo, A.; Matsumura, S.; Mitsunori, Y.; Ochiai, T.; Tanaka, H.; Tanabe, M. Distinct clinicopathological phenotype of hepatocellular carcinoma with ethoxybenzyl-magnetic resonance imaging hyperintensity: Association with gene expression signature. *Am. J. Surg.* **2015**, *210*, 561–569. [[CrossRef](#)] [[PubMed](#)]
152. McCann, S.M.; Jiang, Y.L.; Fan, X.B.; Wang, J.N.; Antic, T.; Prior, F.; VanderWeele, D.; Oto, A. Quantitative multiparametric MRI features and pten expression of peripheral zone prostate cancer: A pilot study. *Am. J. Roentgenol.* **2016**, *206*, 559–565. [[CrossRef](#)] [[PubMed](#)]
153. Stoyanova, R.; Pollack, A.; Takhar, M.; Lynne, C.; Parra, N.; Lam, L.L.C.; Alshalalfa, M.; Buerki, C.; Castillo, R.; Jorda, M.; et al. Association of multiparametric MRI quantitative imaging features with prostate cancer gene expression in MRI-targeted prostate biopsies. *Oncotarget* **2016**, *7*, 53362–53376. [[CrossRef](#)] [[PubMed](#)]
154. Untch, M.; Gerber, B.; Harbeck, N.; Jackisch, C.; Marschner, N.; Mobus, V.; von Minckwitz, G.; Loibl, S.; Beckmann, M.W.; Blohmer, J.U.; et al. 13th St. Gallen international breast cancer conference 2013: Primary therapy of early breast cancer evidence, controversies, consensus—Opinion of a german team of experts (Zurich 2013). *Breast Care* **2013**, *8*, 221–229. [[CrossRef](#)] [[PubMed](#)]
155. Mirabelli, P.; Inconato, M.; Coppola, L.; Infante, T.; Parente, C.A.; Nicolai, E.; Soricelli, A.; Salvatore, M. SDN biobank: Bioresource of human samples associated with functional and/or morphological bioimaging results for the study of oncological, cardiological, neurological, and metabolic diseases. *Open J. Bioresour.* **2017**, *4*. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).