Grade IV glioblastoma multiforme (GBM) are the most aggressive brain tumors, with an extremely poor prognosis, dismal median overall survival rate of approximately 12–15 months with standard treatment, and a relative 2-year survival of only 30% (1). Lower-grade gliomas (grades I, II and III) comprise the remainder of primary malignant gliomas, and have a relative 2-year survival rate of 70–90% (2). Approximately 17,000 new cases of malignant gliomas are diagnosed each
year in children, adolescents and adults, at an occurrence rate of about 5 in 100,000. Of these cases, 60–70% are GBM, which result in a mortality rate of over 10,000 deaths each year (3). Some GBM tumors manifest as primary tumors, and others show signs of progression from a lower-grade glioma (4). The aggressiveness of GBM has driven development of new surgical techniques, anti-angiogenic therapies, immunotherapies, and improved radiotherapies. The response to treatment has been correlated with molecular classification and subtyping based on genetics and expression data (5-7). Despite these efforts to better understand the underlying biology of GBM and advancements in the clinical treatment of this disease, it remains one of the most recalcitrant tumor types.

Like other solid tumor types, GBM develops a heterogeneous pattern of mutations (5). GBM does not feature sequential characteristic driver mutations, with multiple alterations occurring early in tumor pathogenesis, making the development of targeted therapies particularly challenging (8). Individual mutations or chromosomal alterations have not been linked with stages in tumor progression, unlike colorectal (9) and prostate cancers (10). In contrast, progression of colorectal cancer (CRC) has been associated with step-wise mutations and chromosomal alterations (9). Similarly, evidence from mouse models of prostate cancer as well as clinical observations indicate that sequential alterations in p27, NKX3.1, PTEN, and androgen receptor drive the pathogenesis of the disease (10). Women are more likely to have mutations in TP53, since these mutations are more prevalent in secondary GBM, which is more common in women (11), and IDH mutations are more common in adult secondary GBM than in children (12,13). Recent efforts have shown the predictive and prognostic utility of genetic characterization of GBM (14,15). Mutation, copy number, and expression data have been used to segregate GBM into four genomically-defined subtypes: classic, mesenchymal, neural, and proneural (7). Mutation and expression data can predict patient response to therapy (16), and correlations have been drawn between response to therapy and MGMT promoter methylation (17,18). The molecular heterogeneity represents a major challenge to the development of novel targeted therapies for GBM and evidence-based clinical decision-making (19). High-resolution genetic, epigenetic, and molecular descriptions of the range of GBM phenotypes will likely provide the basis for future improvements in treatment and the development of novel therapies.

**Radiomics workflow**

As discussed above, solid tumors have heterogeneous mutations, copy number alterations, and chromosomal aberrations across the tumor volume (20). This intrinsic property has made characterization of tumor phenotypes and the development of targeted therapies particularly challenging (21). Given the heterogeneity of GBM and other cancer types, numerous imaging approaches have been taken to comprehensively characterize tumors (22). In the last decade, radiomics has emerged as the concept of extracting quantitative radiologic features and drawing associations with clinical outcomes in tumors of the breast (23), brain (24), head and neck (25). Imaging data is acquired through the application of a variety of techniques and variants of X-ray computed tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI) (Figure 1A,B). The goal of this approach is to inform clinical decision-making by providing semi-automatically and automatically extracting radiologic features (Figure 1C,D) and associating these factors with outcomes like progression and survival. Clinical and biological associations are made through data mining, hypothesis generation, and biomarker discovery (Figure 1E). Given the complex intra-tumoral and inter-patient heterogeneity characteristic of GBM, and the difficulty in obtaining representative biopsies from which detailed molecular information can be extracted (26), sophisticated imaging approaches have the potential to address the tumor heterogeneity problem (27). It is hypothesized that it is possible to extract detailed phenotypic information by processing radiological imaging data (22). Determining tumor genetics and expression patterns from radiologic features, and developing these features as prognostic and predictive markers is an exciting possibility (28). Additionally, radiomics analysis has the potential to distinguish those low-grade gliomas which will progress to GBM from those which will not by determining the underlying genetic and molecular indicators of progression (4). Thus radiomics provides an avenue to tackle the formidable challenges of cancer treatment based on image-derived appearance, especially in the case of GBM.

In the past, radiology was analyzed qualitatively, following clinical algorithms to determine patient response to therapy and disease progression (29). Standardized measures of tumor volume by MRI, CT and ultrasound were incorporated into the RECIST criteria for tumor...
response (30), and advances in the use of PET modalities have been used in the more recently published PERCIST tumor response criteria (31). Developments in image analysis have allowed quantitative information to be derived from medical imaging. One-dimensional histogram-based and two-dimensional co-occurrence texture analyses (32) were developed to study and compare MRI (33) and other diagnostic images (34). Texture analysis has been used to predict clinical responses in non-small cell lung cancer patients (35) and was shown to be capable of discriminating between prostate tumor Gleason scores (36). This approach has been taken further by making associations between GBM tumor morphology as seen in MRI and underlying genetics (37) and expression data (38). Analysis of GBM tumor MRI imaging revealed novel imaging biomarkers capable of predicting clinical outcomes (39). These studies illustrate the potential of imaging data to overcome the limitations of traditional biomarkers.

Strategies—qualitative and quantitative measurement methods

Two main approaches have been taken to develop features for radiomic studies in GBM. The first has been to create standardized semantic features (Figure 2) which can be reproducibly scored by radiologists. This data is generated manually or by semi-automated methods. The second approach has been to derive fully computational features using imaging and statistical techniques. Both semantic and computational radiomic features are derived from multiple imaging techniques and modalities, including MR, PET, and CT. Semantic and computational features must take consideration of their dependence on scanning and acquisition protocols, signal-to-noise ratio and image resolution variations, properties unique to each modality and technique.

Several studies have correlated semantic features with clinical outcomes. Necrosis and tumor enhancement were...
Some common VASARI (visually accessible rembrandt images) features –
- Tumor Location
- Lesion Size
- Proportion Enhancing
- Proportion Necrosis
- Proportion Edema
- Definition of the Enhancing Margin

**Figure 2** Illustration of semantic features.

**Figure 3** Illustration of computational features extracted from MRI. (A) Spatially-derived habitats from the imaging modalities to define the characteristics of the tumor region; (B) texture features derived from haralick computations to compare tumor characteristics such as homogeneity, entropy, correlation, etc. Below, two types of Haralick features, namely—correlation (measure of homogeneity) and energy (measure of angular moment) are computed on two different GBM patients. MRI, magnetic resonance imaging; GBM, glioblastoma multiforme.

Identified as prognostic indicators (40). Non-contrast-enhancing tumor, multifocality, necrosis, satellites, and edema correlated with prognosis and survival (41). Iterative scatter search combined with an induction learning algorithm correlated imaging features with clinical data to predict survival in high-grade gliomas (42). The Visually Accessible Rembrandt Images (VASARI) feature set (Figure 2), comprising thirty semantic features developed to standardize radiological assessment of GBM, predicted survival and molecular subtype (29). These studies indicate the feasibility of developing standardized semantic feature sets and the efficacy of associating semantic features with prognostics and molecular descriptions in GBM.

While the studies outlined above focused on associating semantic features with clinical outcomes and tumor genetics, other studies have derived computational features computed from 2D/3D tumor regions (Figure 3) and related them to the molecular characteristics of GBM. Haralick features...
are first-order statistics which discriminate images based on texture are calculated from a co-occurrence matrix of pixel intensities within a region-of-interest (32). The original set of Haralick texture features was two-dimensional, and the approach has been adapted for segmenting 3D CT data by mathematically re-defining each feature in three dimensions (43). This method was used to accurately segment structures within the abdominal cavity (44). Spatial habitats are tumor-subregions computed across multi-modality imaging sequences [e.g., T1-post contrast, T2 and fluid attenuated inversion recovery (FLAIR) sequences], where each habitat represents a region with a unique combination of “high” and “low” pixel intensities in each imaging sequence, and have been associated with gene expression status of epidermal growth factor receptor and 12-month overall survival status (24) in GBM. Imaging habitats are correlated with GBM molecular subtype status (neural, pro-neural, mesenchymal and classical) and survival status (45). Imaging habitats quantify the grey-level heterogeneity in GBM MR scans (46) and track tumor evolution driven by detecting variations in tumor blood supply (47).

The majority of radiomics studies in GBM have focused on MRI images (27) derived from T1-weighted acquisitions, T2-weighted, and FLAIR (Figure 4). Gadolinium contrast agent is often administered following T1-weighted acquisition. Following this acquisition sequence, GBM generally appears as a non-enhancing region ringed with enhancing signal, with the area surrounding the tumor (peritumoral) being bright in the T2 and FLAIR scans (48,49). The signal acquired by each modality indicates the presence of vasogenic edema, tumor infiltration, peritumoral tissue, etc. (50). T1-T2-T1c-FLAIR modalities were used to develop computational features which correlated strongly with the VASARI semantic feature set (37), and in multiple radiogenomics studies (51) which correlated VASARI features with mRNA expression (38), mutational status (29), dysfunctional metabolism (52), molecular subtype and survival (53). The mesenchymal GBM subtype was identified using T1c-FLAIR images (54). Another study using T1,
T1c, T2 modalities correlated semantic features with epigenetic status in GBM (55). A comprehensive review and tabulation of radiogenomics studies can be found in Bai et al. (51). Diffusion weighted imaging (DWI) and perfusion MR (MRP) variants (Figure 4) provide additional tumor characteristics at the tissue level. Varied combinations of MR modalities have been used to construct radiological feature combinations that highlight distinct portions of the tumor and surrounding tissue. Magnetic resonance spectroscopy imaging (MRSI) proved useful in guiding radiotherapy by defining a molecular signature distinguishing tumor and non-tumor brain tissue (56).

Semantic and computational (57) feature sets have been derived from CT and PET. Like MR, they can be performed using contrast agents or tracer compounds to gain additional physiological information. When CT is performed using a contrast-enhancement, the GBM tumor region presents as an area of low density ringed by an enhancing region (58). PET scans can measure metabolic activity in GBM, but it has not yet been determined whether combined PET/MRI improves diagnostics (59). A number of variants of PET use molecular tracers to derive additional information based on metabolic activity quantified by glucose uptake. FDG PET, which utilizes the fluorodeoxyglucose tracer, has been investigated as an alternative to gadolinium enhancement, but reports conflict as to its efficacy in diagnosing recurrent GBM (60,61). FDG is the most widely used PET tracer (62), and several studies have developed semantic and computational feature sets using FDG-PET (63,64). Computational features from FDG PET/CT images were used to stage lymphoma (65) and non-small cell lung cancer patients (66), and to predict response to radiation therapy in lung cancer (67). In addition to FDG, a large number of alternative tracer compounds have been investigated (62), although there have been few radiomic studies from these less commonly used contrast methods. The FDOPA PET technique, which is based on the DOPA-decarboxylase pathway and amino acid transport, has not been used in radiomics studies to-date, but has been shown to have predictive value in the recurrence of LGG (68,69). Thymidine and FLT PET (70) and amino acid PET have not been studied using the radiomics approach, but these variants of the imaging technique may provide additional prognostic and predictive information. PET PET, which uses a 18-F-fluoroethyltyrosine tracer, has been used to determine the extent of invasion of glioma cells into the surrounding brain matter (71). In PET scans with 18F-FMISO, GBM is characterized by higher uptake values relative to the gray matter at the boundary of the tumor, and the central necrotic tissue of GBM tumors will accumulate less tracer (72,73). Compared to other imaging modalities, PET better differentiates recurrent or residual GBM tumor from edema and scar tissues after resection (74). However, for initial diagnosis of GBM, PET scans in general have low sensitivity and specificity (74). As a result, the availability of pre-treatment GBM PET images is also limited for radiomic analysis.

**Challenges and opportunities**

**Medical Image acquisition and standardization**

Medical image acquisition is routine for standard MR sequences, but acquisition/scanning protocols vary among institutions, leading to challenges in comparing or combining data gathered in multi-center clinical trials. There are several components of the image acquisition process that lead to variation in the data. These include scanner variability, variation in the specifications of an imaging device involved, and procedures followed by a particular radiologist or imaging physicist/technician. The protocol defined by the physicist/radiologist can vary in terms of image resolution, slice thickness, cut angles and washout period for the contrast imaging. Hence, standardization of image acquisition is central to the integrity of the entire radiomics pipeline. If common standards of acquisition cannot be achieved, the imaging pipeline should incorporate methods standardizing the imaging prior to computational/semantic/volumetric feature extraction. There are some solutions in this space, such as intensity normalization, registering multi-parametric data to a specific anatomical plane, isotropic pixel or voxel re-slicing. Standardized image formats allow data to be readily processed across all steps of the radiomics pipeline. The Neuroimaging Informatics Technology Initiative (NIFTI) format provides support for functional MRI, a coordinate system linked to voxel indices, and dual file storage (75). The Digital Imaging and Communications in Medicine (DICOM) engineering standard provides a system facilitating communication between medical imaging devices and software from multiple vendors (76). DICOM objects contain metadata with patient identification, acquisition device information and printing parameters (77). The Nearly Raw Raster Data NRRD format support multiple compression algorithms and data represented as...
Tumor segmentation

Following acquisition, the tumor images can be segmented through a manual, semi-automated or fully-automated procedure. Since there is no defined ground truth, this process will not be perfectly accurate. Consistency of segmentation process plays a significant role in the radiomics pipeline, as variations in processing steps such as thresholding can affect segmentation (90) by altering how the tumor is delineated (91). Fully-automated segmentation pipelines are available for GBM (92), and are under development for LGG. Random forest classifiers were applied to segment GBM tumor volumes for feature extractions (93). Automation improves reproducibility and concordance, since feature generation and the further downstream analysis are dependent on the quality of the initial segmentation. Manual delineation is time-consuming and further increases the chances of inter-observer variability. When manual contouring was compared to a semi-automated approach using 3D-Slicer, the semi-automated approach (94) improved reproducibility and robustness. Pattern recognition software and techniques reduce observer effects by automating the segmentation pipeline, increasing the robustness of derived results.

Feature Interpretation and analysis

High-throughput feature extraction is at the core of the radiomics process. As described above, there are two types of features—semantic and computational. The definition of these features is dependent on the hypothesis of the project. The numerous methods to obtain the features result in thousands of complex descriptions of the region of interest. Moving from manual to automated methods of feature extraction, automated feature extraction complements manual analysis and reduces variation in scoring semantic features. However, automated feature extraction is still vulnerable to site-specific variations in image acquisition, and any automated method may require modifications when implemented at a study site. Due to different methodologies, reproducibility and robustness of these extracted features is vital as these features will directly determine the correlation drawn to tumor genomic, expression, and microenvironment phenotypes. For example, differences in thresholding stringency can alter the number and attributes of extracted features (95). Outcome modeling based on computational features is a complex, multi-step process starting with pre-processing, followed by feature estimation, feature selection, classification, and finally evaluation by validation studies (96). Robustness of features can be analyzed on the basis of factors such as geometric transformations of the regions of interest and intensity variability (97). We analyzed the robustness of texture features through 8 different geometric transformations of ROIs (horizontal translation by 2 pixels, horizontal and vertical translation...
by 2 pixels, rotation by 1-degree, rotation by 5-degree, moving each point on the outline on the horizontal and vertical axes by a zero-median random number with a 0.1 and 0.5 pixel standard deviation, shrinking the ROI by 1 pixel and dilating the ROI by 1 pixel) (37). Eighty-two TCGA GBM cases were used, using the original as well as the transformed ROI for texture analysis. A set of thirteen Haralick feature ratios with 2 filters [Laplacian of Gaussian (LOG) (98) and Gaussian (99)] at 5 filter widths (0.2, 0.4, 1.5, 2.5, 5) and 4 pixel distances (1, 2, 4 and 8 mm) (32) were computed. Besides this, histogram-based features (uniformity, mean-intensity and entropy) were computed. Additionally, some regional properties such as area, minimum, mean and maximum intensity were computed from the tumor region. All of these features (2,236 for T1-post contrast and 2,236 features for FLAIR) were computed for each of the 9 ROIs (original and 8 transformations) per case in the 82 dataset. Feature extraction was followed by assessment of robustness via intraclass correlation (ICC), with ICC computed between the various features. The correlation cut-off was set as 0.6 and 22 FLAIR and 26 T1-post contrast features were found to be robust. Heatmaps were generated to depict this high correlation. Figure 5 summarizes the results of this study. Tables 1 and 2 show the interpretation of these robust features in the context of the GBM radiomics problem. Image-derived features classify the target hypothesis (outcome). Classification represents the greatest computational challenge, utilizing machine learning algorithms such as neural networks, support vector machines, decision trees, and logistic regression. During data analysis and predictive modeling, it is important to perform cross-validation (96) to evaluate the predictive models, or preferably, to evaluate their performance on a clinically-matched independent (test) cohort. Cross-validation assesses the generalizability of the analysis pipeline for image-biomarker identification. The pipeline can be then used with higher reliability in multicenter/multi-institutional settings. To minimize type I errors caused by multiple comparisons, methods such as the Benjamini-Hochberg correction (102) and bootstrap-based
correction (101) have been proposed. When the sample size is relatively small, methods have been developed to minimize false-positive results (103).

Besides generating automated features on the ROIs in a high-throughput manner, it is important to discuss their reporting structure. A lexicon for radiomics studies, made especially necessary in multi-institutional settings. Improvements have been made in this area through creation of reporting data systems across different disease sites: such as Liver Imaging Reporting and Data System (LI-RADS™, ACR), Prostate Imaging Reporting and Data System (PI-RADS™), Head Injury Imaging Reporting and Data System (HI-RADS, ACR), and Lung Reporting and Data System (LungRADs™, ACR). A universal reporting system with standard terminology and syntax plays a significant role during the feature interpretation and data analysis (104).

### Creating mineable data

While performing classification with high-dimensional radiomics data (as might be obtained by large-scale feature extraction), there is potential risk of overfitting to the data. The risk of overfitting increases when the number of instances is much fewer than the number

---

**Table 1** Interpretation of robust FLAIR features

<table>
<thead>
<tr>
<th>FLAIR features</th>
<th>Type [regional, TxRAD (100), Haralick (32)]</th>
<th>Ratio</th>
<th>Filter [Gauss (99) and LOG (98)]</th>
<th>Sigma</th>
<th>Distance</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1978 TxRAD</td>
<td>Ratio 2</td>
<td>Gauss</td>
<td>0.2</td>
<td></td>
<td></td>
<td>Histogram uniformity</td>
</tr>
<tr>
<td>V54 Regional</td>
<td>Not a ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Regional property-min intensity</td>
</tr>
<tr>
<td>V1821 TxRAD</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>1.5</td>
<td>1</td>
<td></td>
<td>Mean intensity histogram</td>
</tr>
<tr>
<td>V886 TxRAD</td>
<td>Ratio 1</td>
<td>Gauss</td>
<td>5</td>
<td>1</td>
<td></td>
<td>Histogram entropy</td>
</tr>
<tr>
<td>V176 Haralick</td>
<td>Ratio 1</td>
<td>LOG</td>
<td>0.4</td>
<td>2</td>
<td></td>
<td>Sum variance</td>
</tr>
<tr>
<td>V2081 TxRAD</td>
<td>Ratio 2</td>
<td>Gauss</td>
<td>1.5</td>
<td></td>
<td></td>
<td>Mean intensity histogram</td>
</tr>
<tr>
<td>V1267 Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td>2</td>
<td></td>
<td>Sum average</td>
</tr>
<tr>
<td>V1268 Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td>2</td>
<td></td>
<td>Sum variance</td>
</tr>
<tr>
<td>V53 Regional</td>
<td>Not a ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Regional property-area</td>
</tr>
<tr>
<td>V1719 TxRAD</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.2</td>
<td></td>
<td></td>
<td>Histogram uniformity</td>
</tr>
<tr>
<td>V1717 TxRAD</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.2</td>
<td></td>
<td></td>
<td>Mean intensity histogram</td>
</tr>
<tr>
<td>V2186 TxRAD</td>
<td>Ratio 2</td>
<td>Gauss</td>
<td>5</td>
<td></td>
<td></td>
<td>Histogram entropy</td>
</tr>
<tr>
<td>V1769 TxRAD</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td></td>
<td></td>
<td>Mean intensity histogram</td>
</tr>
<tr>
<td>V1875 TxRAD</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>2.5</td>
<td></td>
<td></td>
<td>Histogram uniformity</td>
</tr>
<tr>
<td>V1242 Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.2</td>
<td>8</td>
<td></td>
<td>Sum variance</td>
</tr>
<tr>
<td>V1266 Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td>2</td>
<td></td>
<td>Inverse difference moment</td>
</tr>
<tr>
<td>V1349 Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>1.5</td>
<td>8</td>
<td></td>
<td>Difference variance</td>
</tr>
<tr>
<td>V1207 Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.2</td>
<td>1</td>
<td></td>
<td>Difference entropy</td>
</tr>
<tr>
<td>V1243 Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.2</td>
<td>8</td>
<td></td>
<td>Sum entropy</td>
</tr>
<tr>
<td>V337 Haralick</td>
<td>Ratio 1</td>
<td>LOG</td>
<td>5</td>
<td>2</td>
<td></td>
<td>Cluster shade</td>
</tr>
<tr>
<td>V1271 Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td>2</td>
<td></td>
<td>Difference variance</td>
</tr>
<tr>
<td>V1411 Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>5</td>
<td>1</td>
<td></td>
<td>Sum variance</td>
</tr>
</tbody>
</table>

FLAIR, fluid attenuated inversion recovery; Gauss, Gaussian filter; LOG, Laplacian of Gaussian filter.
of computationally-extracted features. This “curse-of-dimensionality” can be addressed by minimizing the number of features using principal components analysis (PCA) (105), sparse PCA (sPCA), partial least squares regression (PLS) (106), non-linear PCA (107), and auto-encoders (108), in addition to cross-validation for generalized model construction. Some of these feature selection methods can be supervised (using the label information from each instance), or unsupervised (based on exploiting variance in the data). Such dimension reduction methods make interpretation of the reduced feature set more difficult as features are mathematically-combined into composite features.Aside from relating radiomics (imaging-phenotype) with genetic characteristics, there is also a need for modeling formalisms that integrate measurements across these diverse modalities to drive decision-making in

<table>
<thead>
<tr>
<th>T1 features</th>
<th>Type [regional, TxRAD (101), Haralick (32,102)]</th>
<th>Ratio</th>
<th>Filter [Gauss (102) and LOG (100)]</th>
<th>Sigma</th>
<th>Distance</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1789</td>
<td>TxRAD</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td></td>
<td>Mean intensity histogram</td>
</tr>
<tr>
<td>V1272</td>
<td>Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td>2</td>
<td>Difference entropy</td>
</tr>
<tr>
<td>V1977</td>
<td>TxRAD</td>
<td>Ratio 2</td>
<td>Gauss</td>
<td>0.2</td>
<td></td>
<td>Mean intensity histogram</td>
</tr>
<tr>
<td>V1267</td>
<td>Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td>2</td>
<td>Sum variance</td>
</tr>
<tr>
<td>V176</td>
<td>Haralick</td>
<td>Ratio 1</td>
<td>LOG</td>
<td>0.4</td>
<td>2</td>
<td>Sum variance</td>
</tr>
<tr>
<td>V1285</td>
<td>Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td>4</td>
<td>Difference entropy</td>
</tr>
<tr>
<td>V1256</td>
<td>Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td>1</td>
<td>Sum entropy</td>
</tr>
<tr>
<td>V1927</td>
<td>TxRAD</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>5</td>
<td></td>
<td>Histogram uniformity</td>
</tr>
<tr>
<td>V144</td>
<td>Haralick</td>
<td>Ratio 1</td>
<td>LOG</td>
<td>0.2</td>
<td>8</td>
<td>Energy</td>
</tr>
<tr>
<td>V2031</td>
<td>TxRAD</td>
<td>Ratio 2</td>
<td>Gauss</td>
<td>0.4</td>
<td></td>
<td>Histogram uniformity</td>
</tr>
<tr>
<td>V2083</td>
<td>TxRAD</td>
<td>Ratio 2</td>
<td>Gauss</td>
<td>1.5</td>
<td></td>
<td>Histogram uniformity</td>
</tr>
<tr>
<td>V175</td>
<td>Haralick</td>
<td>Ratio 1</td>
<td>LOG</td>
<td>0.4</td>
<td>2</td>
<td>Sum average</td>
</tr>
<tr>
<td>V1200</td>
<td>Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.2</td>
<td>1</td>
<td>Sum of variance</td>
</tr>
<tr>
<td>V54</td>
<td>Regional</td>
<td>Not a ratio</td>
<td>Regional property-min intensity</td>
<td></td>
<td></td>
<td>Regional property-min intensity</td>
</tr>
<tr>
<td>V1449</td>
<td>Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>5</td>
<td>8</td>
<td>Sum average</td>
</tr>
<tr>
<td>V679</td>
<td>TxRAD</td>
<td>Ratio 1</td>
<td>Gauss</td>
<td>0.2</td>
<td></td>
<td>Histogram uniformity</td>
</tr>
<tr>
<td>V2029</td>
<td>TxRAD</td>
<td>Ratio 2</td>
<td>Gauss</td>
<td>0.4</td>
<td></td>
<td>Mean intensity histogram</td>
</tr>
<tr>
<td>V2081</td>
<td>TxRAD</td>
<td>Ratio 2</td>
<td>Gauss</td>
<td>1.5</td>
<td></td>
<td>Mean intensity histogram</td>
</tr>
<tr>
<td>V1242</td>
<td>Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.2</td>
<td>8</td>
<td>Sum variance</td>
</tr>
<tr>
<td>V53</td>
<td>Regional</td>
<td>Not a ratio</td>
<td>Regional property-area</td>
<td></td>
<td></td>
<td>Regional property-area</td>
</tr>
<tr>
<td>V1244</td>
<td>Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.2</td>
<td>8</td>
<td>Entropy</td>
</tr>
<tr>
<td>V1254</td>
<td>Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td>1</td>
<td>Sum average</td>
</tr>
<tr>
<td>V1197</td>
<td>Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.2</td>
<td>1</td>
<td>Energy</td>
</tr>
<tr>
<td>V1226</td>
<td>Haralick</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.2</td>
<td>4</td>
<td>Sum of variance</td>
</tr>
<tr>
<td>V1771</td>
<td>TxRAD</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>0.4</td>
<td></td>
<td>Histogram uniformity</td>
</tr>
<tr>
<td>V1925</td>
<td>TxRAD</td>
<td>Ratio 2</td>
<td>LOG</td>
<td>5</td>
<td></td>
<td>Mean intensity histogram</td>
</tr>
</tbody>
</table>

FLAIR, fluid attenuated inversion recovery; Gauss, Gaussian filter; LOG, Laplacian of Gaussian filter.
the clinical realm.

Conclusions

Radiomics promises to improve the characterization of radiological datasets and provide further insight to guide patient care in the era of personalized medicine. GBM is one of the most genetically heterogeneous tumor types, exhibiting remarkable inter- and intra-patient variability. Substantial progress has already been made in solving many of the technical hurdles inherent in the radiomics process in GBM. Advances in genome sequencing, expression profiling and machine learning have increased the resolution of datasets and the sensitivity and specificity of the computational methods used to analyze them. Statistical models are needed which relate imaging features to GBM molecular status with high specificity/sensitivity to make the approach useful in practice. More studies correlating radiomic features with disease outcomes and molecular attributes will illuminate the underlying tumor biology of imaging features and treatment responses. Large-scale decision algorithms that fuse features obtained across imaging, genomic and clinical modalities can enable multi-modal decision making in the personalized medicine arena.

Acknowledgements

Funding: A. Rao was supported by CCSG Bioinformatics Shared Resource P30 (CA01667), an Institutional Research Grant from The University of Texas MD Anderson Cancer Center (MD Anderson) and a Career Development Award from the MD Anderson Brain Tumor SPORE.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References


67. Lovinfosse P, Janvary ZL, Coucke P, et al. FDG PET/CT
84. Li G, Miller RW. Volumetric Image Registration of Multi-modality Images of CT, MRI and PET. Bethesda, Maryland, USA: National Cancer Institute, National Institutes of Health: 2010.
94. Parmar C, Rios Velazquez E, Leijenaar R, et al. Robust...


