Identifying the mesenchymal molecular subtype of glioblastoma using quantitative volumetric analysis of anatomic magnetic resonance images

Kourosh M. Naeini, Whitney B. Pope, Timothy F. Cloughesy, Robert J. Harris, Albert Lai, Ascia Eskin, Reshmi Chowdhury, Heidi S. Phillips, Phioanh L. Nghiemphu, Yalda Behbahanian, and Benjamin M. Ellingson

Department of Radiological Sciences (K.M.N., W.B.P., R.J.H., Y.B., B.M.E.); Department of Biomedical Physics (R.J.H., B.M.E.); Department of Biomedical Engineering (B.M.E.); Department of Neurology (T.F.C., A.L., R.C., P.L.N.); Department of Human Genetics, David Geffen School of Medicine, University of California–Los Angeles, Los Angeles, California (A.E.); Department of Tumor Biology and Angiogenesis, Genentech, Inc., San Francisco, California (H.S.P.)

Background. Subtypes of glioblastoma multiforme (GBM) based on genetic and molecular alterations are thought to cause alterations in anatomic MRI owing to downstream biological changes, such as edema production, blood–brain barrier breakdown, and necrosis. The purpose of the current study was to identify a potential relationship between imaging features and the mesenchymal (MES) GBM subtype, which has the worst patient prognosis.

Methods. MRIs from 46 patients with histologically confirmed GBM were retrospectively analyzed. The volume of contrast enhancement, regions of central necrosis, and hyperintensity of T2/FLAIR were measured. Additionally, the ratio of T2/FLAIR hyperintense volume to the volume of contrast enhancement and necrosis was calculated.

Results. The volume of contrast enhancement, volume of central necrosis, combined volume of contrast enhancement and central necrosis, and the ratio of T2/FLAIR to contrast enhancement and necrosis were significantly different in MES compared with non-MES GBM (Mann–Whitney, \( P < .05 \)). Receiver-operator characteristics indicated that these 4 metrics were all significant predictors of the MES phenotype. The volume ratio of T2 hyperintensity to contrast enhancement and central necrosis was significantly lower in MES vs non-MES GBM (\( P < .0001 \)), was a significant predictor of the MES phenotype (area under the curve = 0.93, \( P < .001 \)), and could be used to stratify short- and long-term overall survival (log-rank, \( P = .0064 \) using cutoff of 3.0). These trends were also present when excluding isocitrate dehydrogenase 1 mutant tumors and incorporating covariates such as age and KPS score.

Conclusions. Results suggest that volume ratio may be a simple, cost-effective, and noninvasive biomarker for quickly identifying MES GBM.

Keywords: GBM, glioblastoma, IDH1, mesenchymal, molecular subtypes, MRI, radiogenomics.

Glioblastoma multiforme (GBM) is the most common form of malignant glioma, characterized by genetic instability, intratumoral histopathological variability, and relatively unpredictable clinical behavior. According to a recent study by The Cancer Genome Atlas Research Network, GBM should not be considered a single disease but rather should be categorized by molecular subtypes, each with a different sensitivity to therapy. For example, patients whose tumors have a signature enriched in genes associated with neural development (proneural [PN])
have been shown to have better survival compared with those who have signatures resembling the mesenchyme (mesenchymal [MES]), potentially related to increased treatment resistance in the latter. Based on the significantly different prognoses in MES compared with the other phenotypes, we hypothesized that an imaging signature derived from standard preoperative MRIs may be able to noninvasively identify GBM with the MES signature.

Radiogenomics is a relatively new discipline that aims to establish empirical and biological relationships between radiographic imaging features and “-omic” signatures, including morphometric, genomic, molecular, and proteomic characteristics. Several links between well-known radiographic features of tumors and “-omic” signatures have been established in a variety of cancer types, including hepatocellular carcinoma, liver cancer, lung cancer, and GBM. Mutant isocitrate dehydrogenase (IDH) which are associated with secondary GBM and a favorable prognosis, have been shown to be associated primarily with the PN subtype. Additionally, links have been established among the amount of contrast enhancement, tumor location, and IDH1 mutational status. Based on these relationships, we hypothesized that the volume of T2 hyperintensity and contrast enhancement may provide insight into whether a tumor has the MES molecular signature. The current pilot study examined the potential relationship between lesion volume measurements and molecular subtypes of GBM in a total of 46 patients obtained from our institution’s neuro-oncology database.

Methods

Patients

A total of 46 patients with histologically confirmed de novo (primary) GBM and molecular data available were retrospectively examined from our institution’s neuro-oncology database from April 2000 through December 2011. These patients were a subset of a larger cohort used in a previous study examining IDH1 mutation status. Additionally, a total of 43 of the 46 patients in the current study had received upfront (first line) bevacizumab combined with temozolomide and radiotherapy as part of a previous phase II trial. Patient data used in the current study were institutional review board–approved and compliant with the Health Insurance Portability and Accountability Act. Overall survival (OS) was defined from the time of initial diagnosis to death.

Magnetic Resonance Imaging

Presurgical, postcontrast T1-weighted images and either T2-weighted or T2-weighted/fluid attenuated inversion recovery (FLAIR) images were acquired using standard pulse sequences on either 1.5T MR (Siemens Avanto, Siemens Sonata, Siemens Symphony, Siemens Magnetom Vision, Siemens Healthcare; GE Genesis, GE Signa Excite, GE Signa HDx, GE Medical Systems; Philips Interia, Philips Medical Systems) or 3.0T MR (Siemens Trio, Siemens Healthcare). Postcontrast T1-weighted images were acquired after injection of either gadopentetate dimeglumine (Magnevist, Bayer Schering Pharma AG) or gadobenate dimeglumine (Multihance, Bracco Diagnostics), administered at a dose of 0.1 mmol/kg, using an echo time (TE) of 1.5–21 ms, a repetition time (TR) of 5.5–450 ms, and slice thickness of 1.5–5 mm. T2-weighted images were acquired using a TE of 89–183 ms, a TR of 3500–10 000 ms, and slice thickness of 3–5 mm. T2-weighted FLAIR images were acquired using an inversion time of 2200 ms, a TE of 90–150 ms, a TR of 5000–10 000 ms, and slice thickness of 3–5 mm.

Regions of Interest

Three regions of interest were examined using custom scripts in Analysis of Functional NeuroImages software (http://afni.nimh.nih.gov/afni): (i) contrast enhancement (hyperintensity), (ii) central necrosis (hypointensity) on postcontrast T1-weighted images, and (iii) T2 hyperintensity (excluding necrosis and contrast enhancement) on either T2-weighted or T2-weighted/FLAIR images (Fig. 1). The investigator who performed the volumetric analysis (K.N.) and 2 investigators who independently verified the contours (W.B.P. and B.M.E.) were all blinded to the phenotype results until completion of the study. Additionally, the ratio of T2 hyperintense lesion volume to contrast enhancement plus necrosis volume [FLAIR/(enhancement + necrosis)] was calculated and explored as an additional biomarker for predicting GBM subtypes.

Molecular Subtypes

Gene expression microarray analysis was performed using standard, previously published preparation and analysis protocols. Gene expression subclassification was performed using the Hierarchical Clustering (HC) classification determined via the gene voting strategy established by Freije et al and specific subclasses defined by Phillips et al. Briefly, the mean value of each probe set was evaluated from all samples within the specific microarray platform. Then, probe sets from each sample were assigned a “yes” or “no” vote if that probe set’s value was above or below the probe set mean. Next, the yes or no votes for the probe sets were tallied and used to categorize every GBM into 1 of 3 HC molecular subtypes. Subtype names were chosen based on the expression of signature genes: PN, proliferative (PROLIF), and MES. The PN subtype typically has histological markers of Olig2, DLL3, and BCAN; lacks any particular chromosome gain or loss; and has normal epidermal growth factor receptor (EGFR) and intact phosphatase and tensin homolog (PTEN). The PROLIF phenotype has histological markers for proliferating cell nuclear antigen and...
topoisomerase II alpha, gain of chromosome 7, loss of chromosome 10, PTEN loss, and can have either amplified or normal EGFR. The MES phenotype typically has histological markers for chitinase-3-like protein 1, YKL40, CD44, and vascular endothelial growth factor (VEGF); gain of chromosome 7 and loss of chromosome 10; PTEN loss; and either amplified or normal EGFR expression. Lastly, each tumor was voted into 1 of the HC subtypes based on the highest vote tally (eg, tally of MES probe sets above the mean/total number of MES probe sets). Out of the 46 patients in the current study, 22 had a MES signature, 17 expressed the PN subtype, and 7 were classified as having a PROLIF molecular signature. IDH1 status was also determined, using techniques outlined in a previous publication. A total of 6 of the 46 cases in the current study were determined to be IDH1 mutants.

**Statistical Analysis**

Since neither volume nor volume ratio data followed a normal distribution (Kolmogorov–Smirnov, \( P < .05 \) for all categories), nonparametric statistical tests were used. Specifically, a Kruskal–Wallis test was used to examine differences in volumes and volume ratios between the different GBM subtypes, and Dunn’s test was used for post-hoc comparisons. A Mann–Whitney test was used to compare volumes and volume ratios between the MES phenotype and a group consisting of all non-MES phenotypes. Additionally, receiver-operator characteristic (ROC) analysis was performed for all volumes and volume ratios to determine the sensitivity and specificity for detecting the MES from all non-MES subtypes. Area under the ROC curve (AUC) was used as a measure of ROC performance. Since a total of 5 biomarkers were tested—contrast-enhancing volume, necrotic volume, contrast enhancement + necrotic volume, T2 hyperintense volume, and the ratio of edema/necrosis + contrast—a Bonferroni corrected level of significance of \( P < .01 \) was considered statistically significant. Log-rank analysis on Kaplan–Meier data was performed for comparison of OS between different cox proportional hazards ratio was used for additional multivariate analysis.

**Results**

The volume of T2 or FLAIR hyperintensity, representing both edema and infiltrating tumor, was highest in the PN phenotype (median = 60 cc); however, this was not statistically significant (Kruskal–Wallis, \( P = .4869 \); Fig. 2A), nor was there a statistically significant difference between the MES and non-MES groups after pooling the PN and PROLIF phenotypes (Mann–Whitney test, \( P = .3065 \); Fig. 2B). After Bonferroni correction, a statistically significant difference was observed among the different GBM subtypes when examining the volume of contrast enhancement, excluding central necrosis (Kruskal–Wallis, \( P = .0055 \); Fig. 2C). Specifically, the MES phenotypes, having a median volume of 34 cc, had a significantly higher volume of contrast enhancement compared with the PN and PROLIF subtypes, which each had a median <16 cc (Dunn’s test, \( P < .05 \); MES vs PN and MES vs PROLIF). Similarly, when both PN and PROLIF subtypes were pooled into a non-MES group, the MES phenotype again showed a statistically higher volume of contrast enhancement (Mann–Whitney test, \( P = .0014 \); Fig. 2D). The volume of central necrosis, defined by T1 hypointensity on post-contrast T1-weighted images, was not significantly different among subtypes, after taking into consideration Bonferroni correction (Kruskal–Wallis, \( P = .0252 \); Fig. 2E); however, the MES phenotype had a significantly higher volume of necrosis compared with the PN and PROLIF subtypes after they were pooled into a single group (Mann–Whitney test, \( P = .0071 \); Fig. 2F).

A statistically significant difference was also observed among subtypes when examining the combined volume of contrast enhancement and necrosis (Kruskal–Wallis, \( P = .0012 \); Fig. 3B). Gene expression subtypes also differed significantly when combining these volumetric features by calculating the ratio of T2 hyperintense volume to the total contrast-enhancing and necrotic volume (Kruskal–Wallis, \( P < .0001 \); Fig. 3C). Dunn’s test for multiple comparisons suggested a significant difference in this ratio between the MES and both PN and PROLIF phenotypes (Dunn’s test, \( P < .05 \) for MES vs PN and MES vs PROLIF), which was also the case when PN and PROLIF were pooled into a single group (Mann–Whitney, \( P < .0001 \); Fig. 3D).

ROC analysis suggested that T2/FLAIR hyperintense volume was not a significant predictor of MES versus non-MES subtypes (ROC analysis, AUC = 0.59, \( P = .3014 \); Fig. 4A); however, both the volume of contrast enhancement (ROC analysis, AUC = 0.78, \( P = .0013 \); Fig. 4A) and central necrosis (ROC analysis, AUC = 0.73, \( P = .0069 \); Fig. 4A) could differentiate MES from non-MES subtypes with high sensitivity and/or specificity. In particular, a volume of contrast enhancement higher than 22 cc could identify GBM with the MES subtype with 83% sensitivity and 68% specificity, whereas a volume of central necrosis >1.5 cc could identify the MES phenotype with a sensitivity of 46% but a specificity of 91%. Similar to individual features, the combined volume of contrast enhancement and central necrosis could reliably differentiate the MES from non-MES phenotypes (ROC analysis, AUC = 0.78, \( P = .0011 \); Fig. 4B). Specifically, GBM with the MES phenotype was identified with an 80% sensitivity and 64% specificity when tumors were classified as having a combined volume of contrast enhancement and central necrosis exceeding 35 cc. The ratio of T2/FLAIR hyperintense volume to contrast-enhancing volume including central necrosis was a very strong predictor of the MES phenotype, showing a sensitivity of 83% and a specificity of 87% using a cutoff of 2.3, a sensitivity of 100% and specificity of 60% using a threshold...
of 1.0, and a sensitivity of 71% and specificity of 100% using a cutoff of 3.0 (ROC analysis, AUC = 0.93, \( P < .0001 \); Fig. 4C). There were significant differences among the volume of contrast enhancement, the volume of T2/FLAIR hyperintensity, and the volume ratio with respect to ROC performance (1-way ANOVA, \( P = .0149 \); Fig. 4D). Specifically, the volume ratio had a significantly higher AUC compared with the T2/FLAIR hyperintense volume (Tukey’s test, \( P < .05 \)).

As previously documented, MES phenotypes had a significantly shorter OS compared with non-MES tumors (log-rank, \( P = .0026 \); Fig. 4E). Analysis by multivariate Cox proportional hazards ratio further suggested that the MES subtype had significantly shorter OS compared with non-MES subtypes when the gene expression subtypes were combined with age and KPS (Cox regression; MES vs non-MES covariate, \( P = .0038 \); age covariate, \( P = .5182 \); KPS covariate, \( P = .0878 \)). A ratio of T2/FLAIR hyperintense volume to contrast-enhancing and necrosis volume higher than 3.0, which had a sensitivity of 71% and a specificity of 100% for predicting the MES subtype, was also able to stratify patients based on OS (log-rank, \( P = .0064 \); Fig. 4F), where patients with a volume ratio greater than 3.0 were more likely to live longer than...
patients with a volume ratio less than 3.0. Multivariate Cox proportional hazards ratio analysis confirmed the ability of a volume ratio of 3.0 to independently stratify short- and long-term OS when considering age and KPS (Cox regression; ratio 3.0 vs ratio < 3.0 covariate, $P = .0500$; age covariate, $P = .9617$; KPS covariate = 0.1410).

Since IDH1 mutant GBM is known to be primarily of the non-MES subtype\(^1\) and to have very distinct imaging features,\(^2\) we also tested whether the same volumetric differences between MES and non-MES groups would occur when excluding these patients. Similar to the whole population examined in this study ($N = 46$), the combined volume of contrast enhancement and central necrosis was significantly different between MES and non-MES tumors ($\text{Mann–Whitney, } P = .0053$). ROC analysis also confirmed that the combined volume of contrast enhancement and central necrosis was a significant predictor of the MES phenotype ($\text{ROC analysis, AUC} = 0.76$, $P = .0051$), showing a 76% sensitivity and 65% specificity using a volume threshold of 33 cc. T2/FLAIR hyperintense volume was not significantly different between MES and non-MES phenotypes ($\text{Mann–Whitney, } P = .9129$), which was also confirmed with ROC analysis ($\text{AUC} = 0.51$, $P = .9020$). After excluding IDH1 mutant tumors, the volume ratio of T2/FLAIR hyperintensity to contrast enhancement and necrosis was still significantly different.
between MES and non-MES phenotypes (Mann–Whitney, \( P < .0001 \)). ROC analysis confirmed that this ratio was a strong predictor of the MES phenotype (ROC analysis, AUC = 0.92, \( P < .0001 \)), showing an 88% sensitivity and 78% specificity using a threshold of 1.5 and an 82% sensitivity and 87% specificity using a threshold of 2.3. Together, these results suggest that the ratio of T2/FLAIR hyperintense volume to the volume of contrast enhancement and central necrosis may be a simple, yet powerful, biomarker for predicting OS and differentiating MES from non-MES GBM phenotypes, regardless of IDH1 mutation status.

**Discussion**

Molecular subtypes of GBM have different prognoses and potentially different susceptibility to specific treatments. Currently these phenotypes are determined by microarray analysis, which requires a significant amount of tumor tissue obtained at resection or biopsy. Thus noninvasive surrogates for molecular subtypes of GBM could be clinically useful. In this study, we investigated the ability of quantitative volumetric measurements of tumor burden on standard presurgical anatomic MR images to differentiate MES from non-MES GBM phenotypes. Results suggest the volume of contrast enhancement as well as the volume ratio of T2/FLAIR hyperintensity to contrast enhancement to be powerful biomarkers for the MES phenotype.

Previous studies have demonstrated molecular correlates\(^{17}\) of imaging features including multifocality,\(^{18}\) enhancement,\(^{8}\) location,\(^{11,19}\) and edema.\(^{9,20}\) Recently, several molecular subtypes of GBM have been established, based on gene expression data from microarray analysis. The MES phenotype\(^{3}\) is associated with poor prognosis and tends to be present at tumor recurrence regardless of the phenotype at initial presentation. MES GBM tends to have elevated expression levels of VEGF transcripts. As VEGF is a potent mediator of vascular permeability and is induced by hypoxia, it is plausible to hypothesize that MES tumors have higher volumes of contrast enhancement and necrosis compared with the other subtypes. Additionally, IDH1 mutant gliomas tend to be of the PN phenotype, and these tumors typically lack significant contrast enhancement. Consistent with this hypothesis, MES GBM patients in the current study appeared to have a larger volume of contrast enhancement and macroscopic necrosis compared with non-MES phenotypes; however, results

---

**Fig. 3.** Combined volumetric features in GBM molecular subtypes. (A) Combined volume of contrast enhancement and central necrosis in MES, PN, and PROLIF subtypes showing a significantly higher volume in GBM with the MES signature (Kruskal–Wallis, \( P = .0046 \)). (B) Combined volume of contrast enhancement and central necrosis for MES and non-MES subtypes (Kruskal–Wallis, \( P = .0012 \)). (C) Ratio of T2/FLAIR hyperintense volume to the combined volume of contrast enhancement and central necrosis in MES, PN, and PROLIF subtypes (Kruskal–Wallis, \( P < .0001 \); Dunn’s test for multiple comparisons, \( P < .05 \) for MES vs PN and MES vs PROLIF). (D) Ratio of T2/FLAIR hyperintense volume to the combined volume of contrast enhancement and central necrosis in MES and non-MES subtypes (Mann–Whitney, \( P < .0001 \)).
from the current study suggest that this trend was independent of IDH1 mutation status.

Although not statistically significant, non-MES phenotypes tended to have slightly higher volumes of T2 hyperintensity compared with MES GBM. Since we did not differentiate between non-enhancing tumor burden and vasogenic edema, the reason for this slight difference is not entirely clear. Since MES tumors tend to express more VEGF and tend to be more necrotic, it is conceivable that MES tumors may have more edema compared with non-MES GBM. However, IDH1 tumors tend to be of the PN phenotype and typically lack contrast enhancement but may have a very large non-enhancing component. The lack of statistical differences among phenotypes may therefore be due to the fact that both of these competing processes result in T2 hyperintensity, and we did not selectively differentiate between them. These results

Fig. 4. Receiver-operator characteristics and survival analysis. (A) ROC analysis illustrating that the volume of contrast enhancement and central necrosis is a significant predictor of the MES phenotype (AUC = 0.78, *P* = .0022, and AUC = 0.73; *P* = .0069, respectively), whereas the volume of T2/FLAIR hyperintensity is not (AUC = 0.59, *P* = .3014). (B) ROC analysis of the combined contrast-enhancing and central necrotic volume was a significant predictor of the MES phenotype (AUC = 0.78, *P* = .0011). (C) ROC analysis indicates that the volume ratio of T2/FLAIR hyperintensity to contrast enhancement and necrosis is a significant predictor of the MES phenotype (AUC = 0.93, *P* < .0001). (D) When comparing the AUCs among the 5 biomarkers, results suggested that the volume ratio had significantly better ROC performance compared with T2/FLAIR hyperintense volume (Tukey’s test, *P* = .015). (E) Log-rank analysis of Kaplan–Meier data indicated a significant OS advantage for non-MES compared with MES GBM (log-rank, *P* = .0026). (F) A volume ratio of 3.0, corresponding to a 71% sensitivity and 100% specificity for predicting the MES phenotype, could also be used to stratify patients by OS (log-rank, *P* = .0064).
appear consistent with a recent study by Zinn et al., which found no significant survival differences when univariate analysis of T2/FLAIR hyperintense volume was performed.

The volume of contrast enhancement was larger and the volume of T2 hyperintensity was smaller in the MES vs non-MES. Specifically, MES tumors tended to have a lower ratio of T2/FLAIR hyperintense volume to volume of contrast enhancement and necrosis compared with non-MES. Using a ratio cutoff of 1.0, this biomarker had a sensitivity of 100% and a specificity of 60%. However, when a ratio cutoff of 2.3 was used, this biomarker had slightly lower sensitivity (83%), but a substantially higher specificity (87%). These results suggest that a simple volume ratio T2 hyperintensity to contrast-enhancing tumor burden may be a powerful biomarker for quickly, cost-effectively, and noninvasively identifying MES from non-MES GBM in the clinic and, by extension, OS.

Although results appear quite robust, it is important to point out a few study limitations that may have influenced our results. Owing to the retrospective nature of the current study, we were unable to standardize image acquisition protocols. Thus, differences in slice thickness, spacing, contrast agent concentration, and scan parameters may have led to measurable errors in our estimates of tumor volume. Additionally, as previously mentioned, we did not differentiate between edema and non-enhancing tumor burden on T2 or FLAIR images. This lack of differentiation could have conceivably reduced our sensitivity to differences between MES and non-MES phenotypes, since MES tumors are likely to have more edema, whereas non-MES tumors may have more non-enhancing tumor burden.

**Conclusion**

We retrospectively analyzed 46 de novo GBM patients with gene array information and defined volumes of contrast enhancement, central necrosis, and T2/FLAIR hyperintensity. The ratio of T2/FLAIR volume to contrast-enhancing and necrotic volume was also calculated. The volume ratio was more effective than any of the other factors in stratifying between MES and non-MES subtypes. This study suggests that the volume ratio can be used as a biomarker for the MES GBM subtype.

**Conflict of interest statement.** None declared.

**Funding**

This work was supported by a National Institutes of Health/National Cancer Institute grant no. R21CA167354 (B.M.E.); a UCLA Institute for Molecular Medicine Seed Grant (B.M.E.); a UCLA Radiology Exploratory Research Grant (B.M.E.); a University of California Cancer Research Coordinating Committee Grant (B.M.E.); an ACRIN Young Investigator Initiative Grant (B.M.E.); Art of the Brain (T.F.C.); the Ziering Family Foundation in memory of Sigi Ziering (T.F.C.); the Singleton Family Foundation (T.F.C.); and the Clarence Klein Fund for Neuro-Oncology (T.F.C.).

**References**


