

## Multiple-response regression analysis links magnetic resonance imaging features to de-regulated protein expression and pathway activity in lower grade glioma

Michael Lehrer<sup>1</sup>, Anindya Bhadra<sup>2</sup>, Visweswaran Ravikumar<sup>1</sup>, James Y. Chen<sup>6,7</sup>, Max Wintermark<sup>3</sup>, Scott N. Hwang<sup>4</sup>, Chad A. Holder<sup>5</sup>, Erich P. Huang<sup>8</sup>, Brenda Fevrier-Sullivan<sup>9</sup>, John B. Freymann<sup>9</sup>, Arvind Rao<sup>1</sup> and TCGA Glioma Phenotype Research Group

<sup>1</sup>Department of Bioinformatics and Computational Biology, University of Texas MD Anderson Cancer Center, Houston, TX, USA

<sup>2</sup>Department of Statistics, Purdue University, West Lafayette, IN, USA

<sup>3</sup>Department of Radiology, Neuroradiology Division, Stanford University, Palo Alto, CA, USA

<sup>4</sup>Diagnostic Imaging, St. Jude Children's Research Hospital, Memphis, TN, USA

<sup>5</sup>Department of Radiology and Imaging Sciences, Division of Neuroradiology, Emory University School of Medicine, Atlanta, GA, USA

<sup>6</sup>University of California San Diego Health System, San Diego, CA, USA

<sup>7</sup>Department of Radiology, San Diego VA Medical Center, San Diego, CA, USA

<sup>8</sup>Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD, USA

<sup>9</sup>Clinical Monitoring Research Program, Leidos Biomedical Research Inc., Frederick National Laboratory for Cancer Research, Frederick, MD, USA

**Correspondence to:** Arvind Rao, **email:** aruppore@mdanderson.org

**Keywords:** imaging-proteomics analysis, radiomics, lower grade glioma, signaling pathway activity, multiple-response regression

**Received:** March 30, 2017

**Accepted:** May 02, 2017

**Published:** June 23, 2017

**Copyright:** Lehrer et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### ABSTRACT

**Background and Purpose:** Lower grade gliomas (LGGs), lesions of WHO grades II and III, comprise 10-15% of primary brain tumors. In this *first-of-a-kind* study, we aim to carry out a radioproteomic characterization of LGGs using proteomics data from the TCGA and imaging data from the TCIA cohorts, to obtain an association between tumor MRI characteristics and protein measurements.

The availability of linked imaging and molecular data permits the assessment of relationships between tumor genomic/proteomic measurements with phenotypic features.

**Materials and Methods:** Multiple-response regression of the image-derived, radiologist scored features with reverse-phase protein array (RPPA) expression levels generated correlation coefficients for each combination of image-feature and protein or phospho-protein in the RPPA dataset. Significantly-associated proteins for VASARI features were analyzed with Ingenuity Pathway Analysis software. Hierarchical clustering of the results of the pathway analysis was used to determine which feature groups were most strongly correlated with pathway activity and cellular functions.

**Results:** The multiple-response regression approach identified multiple proteins associated with each VASARI imaging feature. VASARI features were found to be correlated with expression of IL8, PTEN, PI3K/Akt, Neuregulin, ERK/MAPK, p70S6K and EGF signaling pathways.

## **Conclusion: Radioproteomics analysis might enable an insight into the phenotypic consequences of molecular aberrations in LGGs.**

### **INTRODUCTION**

Lower grade gliomas (LGGs), those lesions graded II and III, comprise 10-15% of primary brain tumors [1]. A fraction of these tumors will progress to grade IV glioma, glioblastoma (GBM) [2, 3]. Treatment of LGGs includes surgical resection, radiotherapy and chemotherapy [4-6]. LGGs are classified into molecular subtypes based on alterations in TP53, isocitrate dehydrogenase (IDH) 1 or 2, telomerase reverse transcriptase (TERT), and the transcriptional regulator ATRX: the mutational status of these genes correlates with clinical responses and outcomes [9]. Despite these findings, challenges remain in the treatment of LGGs.

The complementary assessment of gliomas via imaging and molecular pathology approaches, for understanding the response to treatment and poor outcomes, has spurred the investigation of combining imaging data with molecular (genetic/genomic) information. This approach, termed radiogenomics or imaging-genomics, holds potential to improve LGG/GBM outcomes. Radiogenomics allows inference of molecular characteristics of a tumor from image features obtained via human or computational assessment of pre-operative MR [10] images. This allows study of the heterogeneity of the tumor both spatially and over time [11]. Features of GBM tumor radiology correlate with somatic mutations [12]. In a xenograft mouse model, genetic changes in the tumor resulted in radiologic effects [13]. The radiomics approach has been taken in lung [14], breast [15], prostate [16], as well as brain tumors [17-19], predicting clinical outcomes based on imaging features. Initial correlations were drawn between clinical outcomes and radiological features, and more recent efforts have combined genomic datasets to improve the predictive and prognostic value of imaging biomarkers [20, 21]. More recently, the availability of proteomics datasets permits the study of co-ordinated signaling activity underlying such phenotypic changes. Radiogenomics and Radioproteomics [22] approaches promise to aid characterization of the genotype-phenotype landscape in gliomas, coupled with understanding the molecular underpinnings of tumor-associated image features. Such paradigms also enable a whole tumor assessment of phenotypic heterogeneity as a complement to molecular heterogeneity for the characterization of disease state.

In LGGs however, the radioproteomics approach to associate proteomics measurements with radiological feature sets is yet to be investigated, although a small study of four patients showed correlations between mass spectra and enhancement in the corresponding tumor region [23]. We hypothesized that a regression-based association mining of whole tumor imaging and protein

expression information would identify proteins related to tumor radiological features, tumor morphology (phenotype), and associated signal transduction activity for LGGs, while leveraging protein correlation structure (as might be observed from pathway data). *To our knowledge, this study is the first to integrate radiological features with proteomic measurements in lower grade gliomas.*

### **RESULTS**

Multiple-response regression of the combined imaging and protein expression datasets from 57 patients (demographics given in Table 1) revealed VASARI imaging features significantly associated with protein expression levels. These expression changes were in gene products known to be mutated in LGGs, including EGFR, TP53, IDH, FUBP1, NOTCH1, ATRX, and TERT [17].

#### **VASARI features correlate with differential protein levels**

Based on multiple-response regression, associations (regression coefficients) were computed for each combination of imaging feature/molecule. Molecules (proteins and phospho-proteins) with associations that were statistically significant after correction for multiple comparisons were retained for further analysis (Supplementary Table 2). For example, tumor localization to the frontal lobes was positively correlated with abundance of MYH11, and was negatively associated with eIF4E expression. Each imaging feature was matched in this way with negatively-and positively-correlated lists of molecules from the RPPA dataset.

#### **Ingenuity pathway analysis**

Each VASARI feature was matched with a unique set of associated pathways and biological functions using Ingenuity Pathway Analysis software (Table 2 and Supplementary Table 3). The T1/FLAIR ratio, MRI Necrosis, leptomeningeal reaction, mild enhancement quality, edema, cysts, and tumor localization in the parietal lobe were relatively strongly associated with unique patterns of pathway activation and suppression based on their significant correlation with unique patterns of protein expression (Table 2 and Figure 1B). Clustering analysis of the p-values associated with each VASARI feature and IPA canonical pathway revealed two groups of VASARI features (Figure 1A). Leptomeningeal reaction, edema, mild enhancement quality, cysts, T1/FLAIR ratio, MRI necrosis, and localization of anatomic center of the tumor to the

**Table 1: Patient demographic information**

Mean Age at Diagnosis (Range)	47.2 (22 - 70)
Gender - M/F	24/33
Median Overall Survival (Months)	23.3
Median Disease-Free Survival (Months)	23.3
Oligodendroglioma/Astrocytoma/Oligoastrocytoma	25/13/18
Mean KPS (Range)	89.5 (50 - 100)

Demographics are given for the 57 patients included in this study.

**Table 2: Radiological features are associated with unique pathway alterations in LGG**

T1/FLAIR ratio	Acute Myeloid Leukemia Signaling 6.651, Cancer Drug Resistance by Drug Efflux 5.48, AMPK Signaling 5.38
MRI Necrosis	Cancer Drug Resistance by Drug Efflux 8.968, PI3K/AKT Signaling 8.738 14-3-3-mediated Signaling 8.614
Leptomeningeal reaction	Neuregulin Signaling 15.714, Acute Myeloid Leukemia Signaling 11.297 ErbB Signaling 11.067
Cross-product length	Glucocorticoid Receptor Signaling 3.879, IL-2 Signaling 3.602, ErbB2-ErbB3 Signaling 3.537
Enhancing tumor	Molecular Mechanisms of Cancer 3.505, GADD45 Signaling 2.745, DNA damage-induced 14-3-3 signaling 2.745
Enhancing cortical involvement	ATM Signaling 3.408, Hereditary Breast Cancer Signaling 2.912, ILK Signaling 2.636
Mild enhancement quality	14-3-3-mediated Signaling 6.36, Aryl Hydrocarbon Receptor Signaling 6.231, AMPK Signaling 5.71
Edema	Neuregulin Signaling 15.666, Glioblastoma Multiforme Signaling 14.801, Role of Tissue Factor in Cancer 14.2
Definition of the non-enhancing margin	AMPK Signaling 3.326, Glucocorticoid Receptor Signaling 2.965, Notch signaling 2.144
Definition of the enhancing margin	VEGF Signaling 4.62, PI3K/AKT Signaling 4.379, Molecular Mechanisms of Cancer 4.369
Cysts present	Molecular Mechanisms of Cancer 8.654, Glioblastoma Multiforme Signaling 8.568, AMPK Signaling 8.046
Tumor localization in the parietal lobe	HER-2 Signaling in Breast Cancer 8.177, ErbB Signaling 7.94, Pancreatic Adenocarcinoma Signaling 7.533
Tumor localization to the frontal lobe	UVB-Induced MAPK Signaling 2.205, FLT3 Signaling in Hematopoietic Progenitor Cells 2.095, Role of Tissue Factor in Cancer 1.94

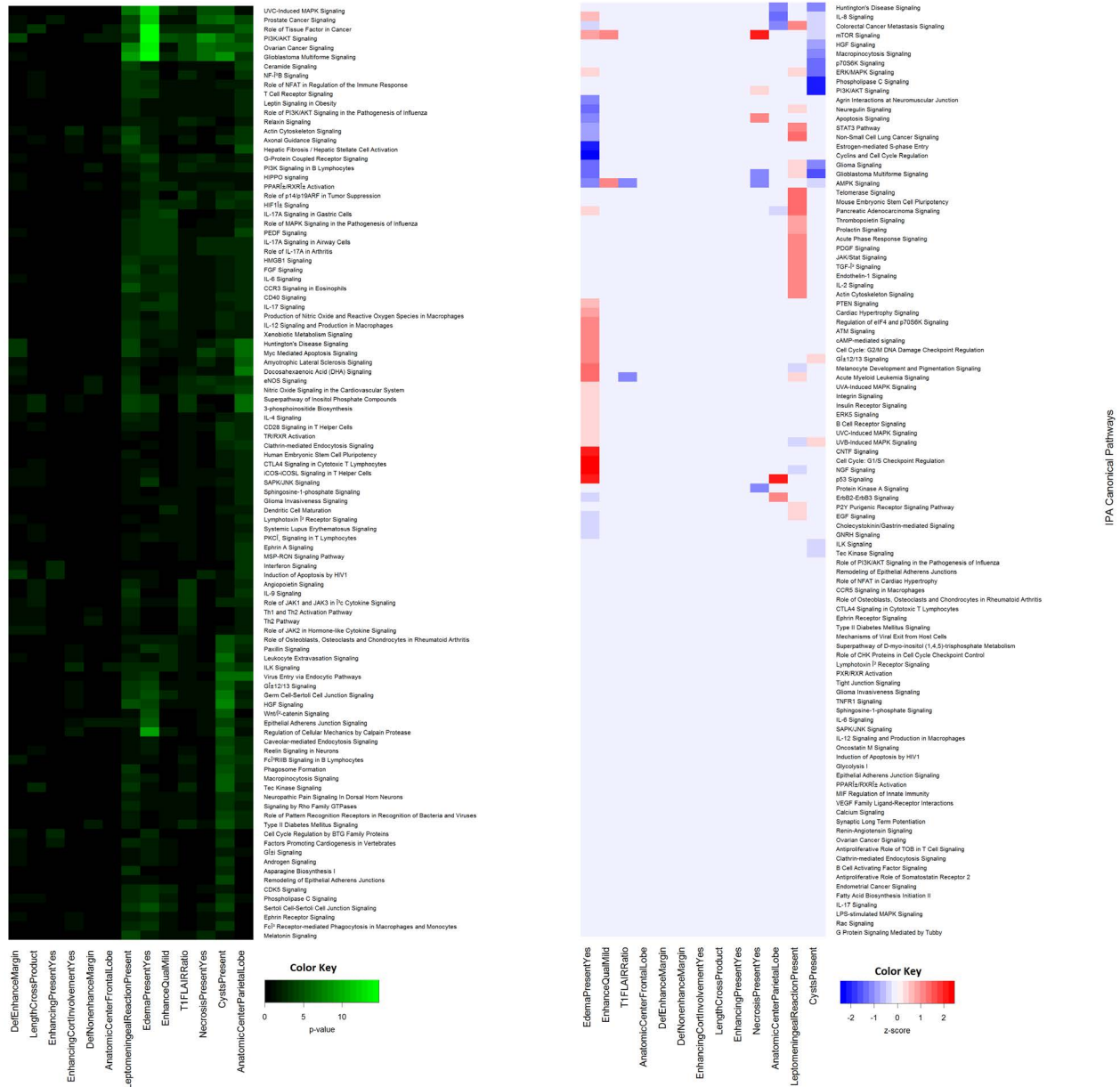
Proteins with expression significantly correlated with imaging features were analyzed by IPA. Top pathways for each feature are shown with the associated  $-\log(p\text{-value})$ .

parietal lobes tended to have lower p-values for each IPA canonical pathway (Figure 1A). Edema, the presence of cysts, T1/FLAIR ratio, Leptomeningeal reaction, localization of tumor to Parietal lobe and MRI necrosis were most strongly associated with altered disease and biological functions (Figure 2A).

From the IPA calculation of Z-activation scores, correlations could be inferred that indicate up regulation and down regulation of a portion of IPA canonical pathways (Figure 1B), diseases, and biological functions (Figure 2B). Focusing on the strongest associations, T1/FLAIR ratio was correlated with down regulation of

AMPK and acute myeloid leukemia signaling. MRI necrosis was associated with up regulation of PI3K/AKT/mTOR signaling and apoptosis, while being correlated with down regulation of AMPK and protein kinase A signaling. Leptomeningeal reaction was associated with up regulation of signaling pathways also important in pancreatic and lung cancers, while being associated with down regulation of NGF and UVB-induced MAPK signaling. Edema was implicated with increased NGF

signaling, as well as increased G1/S checkpoint regulation. The presence of cysts was associated with decreased PI3K/AKT and phospholipase C signaling. Localization of tumor to the parietal lobe was correlated with upregulated p53 signaling activity, and with downregulated IL-8 signaling. These associations highlight the possible functional implications of the correlations drawn between each VASARI feature and protein expression.



**Figure 1: Agglomerative unsupervised hierarchical clustering of Ingenuity Pathway Analysis p-values and Z-activation scores associated with each VASARI features and IPA Canonical Pathways reveal a subset highly-correlated with pathway alterations.** A heat map of a representative subset of the results of the clustering of the p-values (A) and Z-activation scores (B) calculated in IPA for each imaging feature and IPA canonical pathway are shown.





novel diagnostics and prognostics possible [44]. Radioproteomics and radiogenomics show potential to complement proteomics and genomics for understanding the biological underpinnings of phenotypic tumor features and for predicting patient response to treatment [20]. As this approach matures, it is conceivable that imaging could provide strong surrogates to potentially replace traditional physical biopsies, allowing signaling activity status to be determined non-invasively [45].

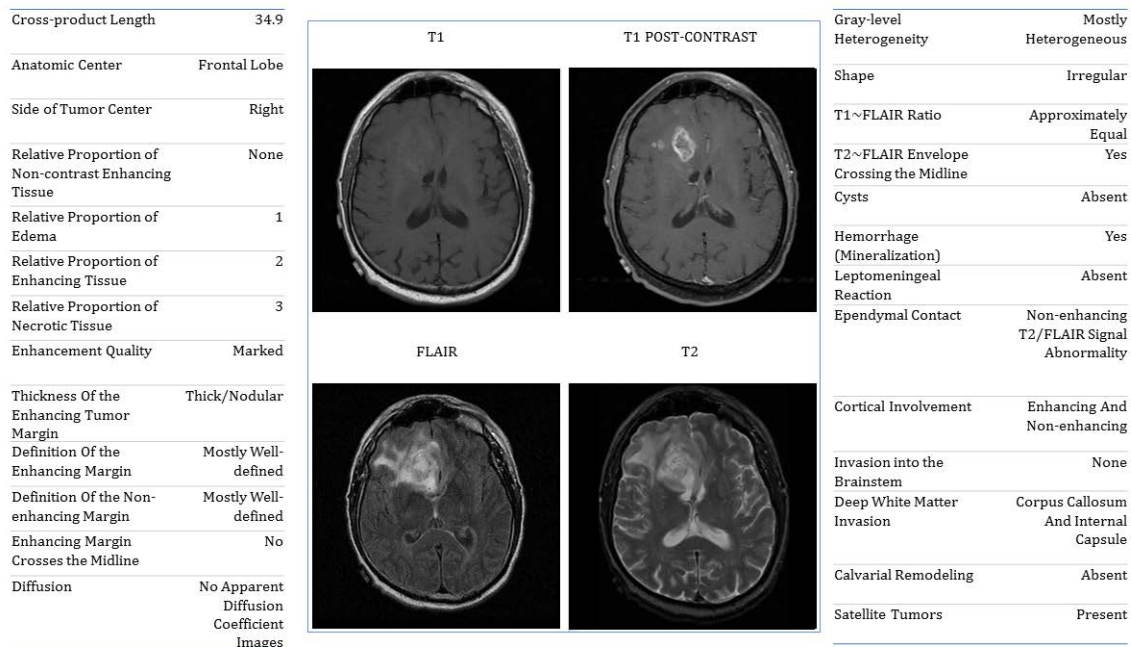
We hypothesized that integrating imaging and protein expression data would reveal imaging biomarkers tying radiological features to the proteomics of LGG. We have demonstrated the use of a novel multiple-response regression approach to integrate imaging features and proteomics data while leveraging the correlation structure between the protein expression values. In the situation where the responses are correlated, and there are multiple predictors associated with multiple responses, standard univariate/multivariate analysis will identify associations incorrectly. Thus the correlated response regression approach is more appropriate in this context, having direct biological significance since it uses the correlation structure of the proteins during regression.

Multiple-response regression of the combined VASARI feature annotations and RPPA molecular expression data revealed significantly correlated proteins and phospho-proteins indicating upregulated signal transduction pathways. Querying the highly-correlated

molecules using Ingenuity Pathway Analysis probed the biological implications of the regression results. Clustering of the pathway analysis highlighted the most important imaging features tied with underlying molecular alterations. By using this analytics pipeline, corresponding biological hypotheses about the underpinnings of these phenotypic characteristics could be within the framework of the IPA knowledge base.

Multiple-response regression revealed associations between imaging features and signal transduction, evidenced by significant correlations with levels of phosphorylated signaling molecules such as Src kinase, Yes-associated protein YAP, ribosomal protein S6, and MAPK, consistent with the established role of these molecules in tumor cell proliferation, migration and invasion. Kinase activity of over-expressed Src drives tumor cell proliferation [43]. YAP is a downstream co-activator in the Hippo pathway, a developmental mechanism frequently co-opted in tumorigenesis and progression [47]. MAPK phosphorylation participates in signaling cell growth and proliferation, and pathways which include these kinases are frequently over-active in tumor cells [48]. Ribosomal protein S6 is a component of the ribosome phosphorylated by S6 kinase, which is a component of the PI3K-mTOR axis, over-activation of which is another common feature of cancer cells [49].

In conclusion, a multiple response regression framework with correlated errors was used to assess the



**Figure 3: Representative axial MR T1, T1 gadolinium contrast enhanced, fluid-attenuation recovery, and T2-weighted acquisitions with corresponding VASARI scores.** Scores for the VASARI feature set for a representative set of MR scans are shown. MR sequences are downloaded from The Cancer Imaging Archive [47].

relationship between image-derived radiologist annotations (features) and protein measurements in TCGA LGG cases. Pathway analysis highlighted the functional effects of proteomic alterations in LGGs on MR imaging features, revealing a unique pattern of protein expression and post-translational modifications associated with each VASARI imaging feature. The downstream effects of these gene expression changes were consistent with the current understanding of tumor biology. One such example is that a physiological event, such as hemorrhage, could cause massive cell death, inducing stress response signaling in affected cells. This was reflected in the results of the pathway analysis, which associated the LGG specific protein expression correlated with hemorrhage with top canonical pathways, including MAP kinase signaling induced by ultraviolet damage. Regression and pathway analysis implicated IL8, PTEN, PI3K/Akt, Neuregulin, ERK/MAPK, p70S6K and EGF signaling pathways, mechanisms that exert control over the cell cycle, growth, and proliferation and are known to be significantly altered in gliomas [17]. These pathway alterations were most significantly associated with a subset of the VASARI features, providing a picture of the molecular underpinnings of the macroscopic radiologic features of LGGs across biological scale.

## MATERIALS AND METHODS

### TCGA low-grade glioma patient dataset

MRI data for 57 patients were obtained from The Cancer Imaging Archive (TCIA), and protein (reverse-phase protein array, RPPA) molecular data were obtained from The Cancer Genome Atlas (TCGA) data portal. These patients correspond to the TCGA/TCIA subset on whom imaging data along with matched proteomics (RPPA) information is jointly available. Clinical data were obtained from cBioPortal [24]. Institutional review board approval was not needed for this retrospective study of public TCGA data.

### MR imaging feature (phenotype) analysis

Adjudicated reader (radiologist) scores of the pre-operative MR images using an updated version of the standardized Visually Accessible Rembrandt Images (VASARI) feature-set were obtained from TCGA Glioma Phenotype Research Group [10]. The TCGA Glioma Phenotype Research Group readers scored cases using an updated version of the original VASARI feature-set, which was developed for assessment of GBMs, modified through expert consensus for lower grade gliomas. The imaging dataset included complete annotations for 27 of the 34 standard VASARI features. The twenty-seven VASARI features included overall shape, tumor length cross product, edema, cysts, hemorrhage (this feature can also result from mineralization), areas of brain invasion,

satellite tumors, enhancement quality, thickness of the enhancing tumor margin, definition of the non-enhancing tumor margin, definition of the contrast-enhancing tumor margin, gray-level heterogeneity, enhancing cortical involvement, leptomeningeal reaction, location of the tumor anatomic center, tumor laterality to right or left cerebral hemisphere, ependymal contact, T1/FLAIR ratio, T2/FLAIR signal crossing the midline, contrast enhancement crossing the midline, calvarial remodeling, and MRI-necrosis (presumed necrosis based upon MR imaging features, as opposed to histology) (Figure 3). As indicated in Zhou et al. [25], the inter-rater analysis via kappa statistic demonstrated good inter-rater agreement along the VASARI features.

### Statistical analysis

To estimate the relationship between the LGG-VASARI imaging features and RPPA-derived protein levels, we used a multiple response regression model developed by Bhadra and Mallick [26]. The model incorporates the VASARI imaging features as the independent variables and the protein expression data (across all the RPPA proteins) as the response vector. The framework aims to capture the associations between the protein expression levels (this is biologically meaningful, since protein expression levels might be correlated), while incorporating the imaging features as covariates in the multiple predictors, multiple responses regression framework (outlined below, more details in the SI). The regression coefficient captures the association strength between the imaging predictor and the protein expression. P-values for the regression coefficients are corrected for multiple testing using Benjamini-Hochberg method [27], and statistical significance is set at 0.05.

### Multiple-response regression with correlated noise: novelty of the proposed framework

Variable selection is a ubiquitous problem in modern genomic data analysis. We provide a brief review below to explain the novelty of our approach more clearly. Further details are in the Supplementary Methods. Bayesian approaches typically perform this through an appropriate sparsity-inducing prior, e.g., spike-and-slab type prior [28, 29], double-exponential prior [30], or the more recent heavy-tailed global-local shrinkage priors such as the horseshoe [31], horseshoe+[32] and generalized double Pareto prior [33]; while frequentist approaches usually work with penalized regression models e.g.,  $L_1$ -norm penalty of the LASSO [33], or combined  $L_1/L_2$  penalty of elastic [34], or combined  $L_1$  and  $L_2$ -norm penalty of Elastic Net [35]. However, in multiple responses regressions, such as the framework considered here, most approaches ignore the correlation structure in the error terms and treat them as independent. It is well-known that ignoring the correlation structure if the true errors are correlated leads to loss of



statistical efficiency, even in low dimensions [36]. Variable selection in presence of correlated errors is being studied only recently. In frequentist settings, Rothman et al. [37] applied an algorithm based on alternating LASSO and graphical LASSO steps to simultaneously select the covariates and the error precision matrix. Yin and Li [38] applied a similar technique to high-dimensional genomic data. Cai et al. [39] designed a two-step procedure for first estimating the residual correlations and then selecting the covariates. We take a Bayesian approach and follow the framework of Bhadra and Mallick [26] in our data analysis. To control for multiple testing correction in the high-dimensional “multiple covariates, multiple responses” framework, we followed a two-stage procedure [26].

1. First we threshold the posterior probabilities of the covariates (imaging features) by controlling FDR at 0.25, yielding a sparse set of imaging predictors.
2. To determine which responses (proteomic measurements) this sparse set of covariates (imaging features) were associated with, we did another level of FDR control. We performed t-tests, of “no association” null hypothesis vs. ‘non-zero association’ alternative hypothesis, between a given imaging feature and a given protein's expression. We then declared those associations significant which had the lowest 10% of the p-values resulting from the above t-tests. All these associations are significant at the adjusted p-value of 0.05. A similar approach has previously been used successfully to analyze eQTL data [26] and glioblastoma data [41].

### **Ingenuity pathway analysis**

Proteins from the RPPA dataset that were found to be strongly-associated with each VASARI feature were queried using the Ingenuity Pathway Analysis software package (IPA™ QIAGEN, Redwood City, CA, <http://www.qiagen.com/ingenuity>). IPA Core Analyses were run on each list of mapped identifiers for each VASARI feature. In the IPA software, p-values were computed by applying the Fisher's exact test based on the number of biological functions, pathways, or molecules in the annotation as defined by the molecules in the selected Reference set, the number of molecules in the Reference set known to be associated with that function, the number of functions, pathways, and molecules in the Reference set, and the number of molecules in the Reference set [42].

### **Hierarchical clustering analysis of imaging-proteomic associations**

To interpret VASARI associations with pathways, diseases and functions, agglomerative unsupervised hierarchical clustering was performed on the p-values associated with each VASARI feature using the “stats”

package in R. Distance matrix calculations were performed using the Euclidean method, and within-cluster variance was minimized using Ward's method [43].

### **Abbreviations**

GBM: Glioblastoma; LGG: Lower grade glioma; WHO: World Health Organization; IPA: Ingenuity Pathway Analysis; VASARI: Visually Accessible Rembrandt Images; TCGA: The Cancer Genome Atlas; TCIA: The Cancer Imaging Archive; RPPA: Reverse Phase Protein Array; IDH: Isocitrate Dehydrogenase; TERT: Telomerase Reverse Transcriptase.

### **Author contributions**

Conception and design: A.R, A.B  
Development of methodology: A.R, A.B, M.L  
Acquisition of data: J.Y.C, M.W, S.N.H, CH, EPH, B.F-S, J.B.F, M.L, V.R, A.R  
Analysis and interpretation of data: M.L, V.R, A.B, A.R,  
Writing, review, and/or revision of manuscript: M.L, A.B, V.R, S.N.H, CH, A.R  
Administrative, technical, or material support (i.e., reporting or organizing data, constructing database): A.B, V.R, J.Y.C, M.W, S.N.H, E.P.H, B.F-S, J.B.F

### **ACKNOWLEDGMENTS**

We acknowledge contributions of Dr. Lee A. D. Cooper, Dr. Adam Flanders, and Dr. C. Carl Jaffe to organizing the analysis of the TCGA LGG imaging and datasets. We also gratefully acknowledge the support of the R and MATLAB communities.

### **CONFLICTS OF INTEREST**

The authors declare no conflicts of interest.

### **FUNDING**

Research in the A.R. laboratory is supported by the NCI Cancer Center Support Grant NCI P30 CA016672, American Cancer Society 627 (RSG-16-005-01), Institutional research grant from The University of Texas MD Anderson Cancer Center, Career Development Award from the Brain Tumor SPORE P50CA127001-07, and Equipment grant from NVidia.

### **REFERENCES**

1. Howlader N, Noone A, Krapcho M. SEER Cancer Statistics Review, 1975-2011, National Cancer Institute. Based on November 2013 SEER data submission, posted to the SEER web site, April 2014. Bethesda, MD. 2014.



2. Ohgaki H, Kleihues P. The definition of primary and secondary glioblastoma. *Clin Cancer Res.* 2013; 19:764–72.
3. Kleihues P, Ohgaki H. Genetics of glioma progression and the definition of primary and secondary glioblastoma. *Brain Pathol.* 1997; 7:1131–36.
4. Le Rhun E, Taillibert S, Chamberlain MC. Current Management of Adult Diffuse Infiltrative Low Grade Gliomas. *Curr Neurol Neurosci Rep.* 2016; 16:15.
5. Yeh SA, Ho JT, Lui CC, Huang YJ, Hsiung CY, Huang EY. Treatment outcomes and prognostic factors in patients with supratentorial low-grade gliomas. *Br J Radiol.* 2005; 78:230–35.
6. Wong ET, Hess KR, Gleason MJ, Jaeckle KA, Kyritsis AP, Prados MD, Levin VA, Yung WK. Outcomes and prognostic factors in recurrent glioma patients enrolled onto phase II clinical trials. *J Clin Oncol.* 1999; 17:2572–78.
7. Field KM, Rosenthal MA, Khasraw M, Sawkins K, Nowak AK. Evolving management of low grade glioma: no consensus amongst treating clinicians. *J Clin Neurosci.* 2016; 23:81–87.
8. Sathornsumetee S, Reardon DA, Desjardins A, Quinn JA, Vredenburgh JJ, Rich JN. Molecularly targeted therapy for malignant glioma. *Cancer.* 2007; 110:13–24.
9. Brat DJ, Verhaak RG, Aldape KD, Yung WK, Salama SR, Cooper LA, Rheinbay E, Miller CR, Vitucci M, Morozova O, Robertson AG, Noushmehr H, Laird PW, et al, and Cancer Genome Atlas Research Network. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *N Engl J Med.* 2015; 372:2481–98.
10. Gutman DA, Cooper LA, Hwang SN, Holder CA, Gao J, Aurora TD, Dunn WD Jr, Scarpace L, Mikkelsen T, Jain R, Wintermark M, Jilwan M, Raghavan P, et al. MR imaging predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set. *Radiology.* 2013; 267:560–69.
11. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJ, Dekker A, Fenstermacher D, Goldof DB, Hall LO, Lambin P, et al. Radiomics: the process and the challenges. *Magn Reson Imaging.* 2012; 30:1234–48.
12. Gutman DA, Dunn WD Jr, Grossmann P, Cooper LA, Holder CA, Ligon KL, Alexander BM, Aerts HJ. Somatic mutations associated with MRI-derived volumetric features in glioblastoma. *Neuroradiology.* 2015; 57:1227–37.
13. Panth KM, Leijenaar RT, Carvalho S, Lieuwe NG, 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; 116:462–66.
14. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebbers F, Rietbergen MM, Leemans CR, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014; 5:4006.
15. Oh J, Sutton E, Veeraraghavan H, Apte A, Morris E, Deasy J. Radiomics and Radiogenomics for Breast Cancer Using Magnetic Resonance Imaging. *Med Phys.* 2015; 42:3588–3588.
16. Stoyanova R, Tachar M, Erho N, Lynne C, Abraham S, Patel M, Buerki C, Lam L, Jorda M, Kryvenko O, Ishkanian A, Abramowitz M, Davicioni E, Pollack A. Radiogenomics of MRI-Guided Prostate Cancer Biopsy Habitats. *Med Phys.* 2015; 42:3605–3605.
17. Brat DJ, Verhaak RG, Aldape KD, Yung WK, Salama SR, Cooper LA, Rheinbay E, Miller CR, Vitucci M, Morozova O, Robertson AG, Noushmehr H, Laird PW, et al, and Cancer Genome Atlas Research Network. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *N Engl J Med.* 2015; 372:2481–98.
18. Yang D, Rao G, Martinez J, Veeraraghavan A, Rao A. Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma. *Med Phys.* 2015; 42:6725–35.
19. Bhat KP, Balasubramanian V, Vaillant B, Ezhilarasan R, Hummelink K, Hollingsworth F, Wani K, Heathcock L, James JD, Goodman LD, Conroy S, Long L, Lelic N, et al. Mesenchymal differentiation mediated by NF-κB promotes radiation resistance in glioblastoma. *Cancer Cell.* 2013; 24:331–46.
20. Nicolasjilwan M, Hu Y, Yan C, Meerzaman D, Holder CA, Gutman D, Jain R, Colen R, Rubin DL, Zinn PO, Hwang SN, Raghavan P, Hammoud DA, et al, and TCGA Glioma Phenotype Research Group. Addition of MR imaging features and genetic biomarkers strengthens glioblastoma survival prediction in TCGA patients. *J Neuroimaging.* 2015; 42:212–21.
21. Colen RR, Vangel M, Wang J, Gutman DA, Hwang SN, Wintermark M, Jain R, Jilwan-Nicolas M, Chen JY, Raghavan P, Holder CA, Rubin D, Huang E, et al, and TCGA Glioma Phenotype Research Group. Imaging genomic mapping of an invasive MRI phenotype predicts patient outcome and metabolic dysfunction: a TCGA glioma phenotype research group project. *BMC Med Genomics.* 2014; 7:30.
22. Djekidel, M. ‘Radiogenomics and radioproteomics.’ *OMICS J. Radiol* 2.2 (2013): 1000115.
23. Hobbs SK, Shi G, Homer R, Harsh G, Atlas SW, Bednarski MD. Magnetic resonance image-guided proteomics of human glioblastoma multiforme. *J Magn Reson Imaging.* 2003; 18:530–36.
24. cBioPortal for cancer genomics. <http://www.cbioportal.org> 2016.

25. Zhou H, Vallières M, Bai HX, Su C, Tang H, Oldridge D, Zhang Z, Xiao B, Liao W, Tao Y, Zhou J, Zhang P, Yang L. MRI features predict survival and molecular markers in diffuse lower-grade gliomas. *Neuro-oncol.* 2017; 19:862–70.
26. Bhadra A, Mallick BK. Joint high-dimensional Bayesian variable and covariance selection with an application to eQTL analysis. *Biometrics.* 2013; 69:447–57.
27. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B.* 1995; 57:289–300.
28. Mitchell TJ, Beauchamp JJ. BAYESIAN VARIABLE SELECTION IN LINEAR-REGRESSION. *J Am Stat Assoc.* 1988; 83:1023–32.
29. George EI, McCulloch RE. VARIABLE SELECTION VIA GIBBS SAMPLING. *J Am Stat Assoc.* 1993; 88:881–89.
30. Park T, Casella G. The Bayesian Lasso. *J Am Stat Assoc.* 2008; 103:681–86.
31. Carvalho CM, Polson NG, Scott JG. The horseshoe estimator for sparse signals. *Biometrika.* 2010; 97:465–80.
32. Bhadra A, Datta J, Polson NG, Willard B. The Horseshoe Estimator of Ultra-Sparse Signals. *arXiv:1502.00560*; 2016.
33. Armagan A, Dunson DB, Lee J. Generalized double pareto shrinkage. *Stat Sin.* 2013; 23:119–43.
34. Tibshirani R. Regression shrinkage and selection via the Lasso. *J R Stat Soc B.* 1996; 58:267–88.
35. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Series B Stat Methodol.* 2005; 67:301–20.
36. Zellner AA. EFFICIENT METHOD OF ESTIMATING SEEMINGLY UNRELATED REGRESSIONS AND TESTS FOR AGGREGATION BIAS. *J Am Stat Assoc.* 1962; 57:348–68.
37. Rothman AJ, Levina E, Zhu J. Sparse Multivariate Regression With Covariance Estimation. *J Comput Graph Stat.* 2010; 19:947–62.
38. Yin J, Li H. A SPARSE CONDITIONAL GAUSSIAN GRAPHICAL MODEL FOR ANALYSIS OF GENETICAL GENOMICS DATA. *Ann Appl Stat.* 2011; 5:2630–50.
39. Cai TT, Li H, Liu W, Xie J. Covariate-adjusted precision matrix estimation with an application in genetical genomics. *Biometrika.* 2013; 100:139–56.
40. Ashraf AB, Daye D, Gavenonis S, Mies C, Feldman M, Rosen M, Kontos D. Identification of intrinsic imaging phenotypes for breast cancer tumors: preliminary associations with gene expression profiles. *Radiology.* 2014; 272:374–84.
41. Bhadra A, Baladandayuthapani V. Ieee. Integrative Sparse Bayesian Analysis of High-dimensional Multi-platform Genomic Data in Glioblastoma. 2013 Ieee International Workshop on Genomic Signal Processing and Statistics (Gensips 2013) 2013:1-4.
42. Krämer A, Green J, Pollard J Jr, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics.* 2014; 30:523–30.
43. Ward J. Hierarchical Grouping To Optimize An Objective Function. *J Am Stat Assoc.* 1963; 58:236.
44. Djekidel M. Radiogenomics and Radioproteomics. *OMICS J Radiol.* 2013; 115.
45. Rosenstein BS, West CM, Bentzen SM, Alsner J, Andreassen CN, Azria D, Barnett GC, Baumann M, Burnet N, Chang-Claude J, Chuang EY, Coles CE, Dekker A, et al, and Radiogenomics Consortium. Radiogenomics: radiobiology enters the era of big data and team science. *Int J Radiat Oncol Biol Phys.* 2014; 89:709–13.
46. Irby RB, Yeatman TJ. Role of Src expression and activation in human cancer. *Oncogene.* 2000; 19:5636–42.
47. Varelas X. The Hippo pathway effectors TAZ and YAP in development, homeostasis and disease. *Development.* 2014; 141:1614–26.
48. Dhillon AS, Hagan S, Rath O, Kolch W. MAP kinase signalling pathways in cancer. *Oncogene.* 2007; 26:3279–90.
49. Magnuson B, Ekim B, Fingar DC. Regulation and function of ribosomal protein S6 kinase (S6K) within mTOR signalling networks. *Biochem J.* 2012; 441:1–21.
50. The Cancer Imaging Archive. <https://publiccancerimagingarchivenet/ncia/> 2016.