UCLA
UCLA Previously Published Works

Title

Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques

Permalink

https://escholarship.org/uc/item/0437n73k

Journal

Neuro-Oncology, 18(3)

ISSN

1522-8517

Authors

Macyszyn, Luke
Akbari, Hamed
Pisapia, Jared M
et al.

Publication Date

2016-03-01

DOI

10.1093/neuonc/nov127

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Neuro-Oncology Advance Access published July 16, 2015

Neuro-Oncology

Neuro-Oncology 2015; **0**, 1–9, doi:10.1093/neuonc/nov127

Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques

Luke Macyszyn†, Hamed Akbari†, Jared M. Pisapia, Xiao Da, Mark Attiah, Vadim Pigrish, Yingtao Bi, Sharmistha Pal, Ramana V. Davuluri, Laura Roccograndi, Nadia Dahmane, Maria Martinez-Lage, George Biros, Ronald L. Wolf, Michel Bilello, Donald M. O'Rourke, and Christos Davatzikos

Department of Neurosurgery, University of Pennsylvania, Philadelphia, Pennsylvania (L.M., J.M.P., M.A., L.R., N.D., D.M.O.); Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania (H.A., X.D., R.L.W., M.B., C.D.); Center for Biomedical Image Computing and Analytics, University of Pennsylvania, Philadelphia, Pennsylvania (H.A., X.D., V.P., M.B., C.D.); Department of Biomedical Informatics, Northwestern University Feinberg School of Medicine, Chicago, Illinois (Y.B., S.P., R.V.D.); Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, Texas (G.B.); Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, Pennsylvania (M.M.-L., D.M.O.)

Corresponding Author: Christos Davatzikos, PhD, Wallace T Miller, Sr. Professor of Radiology, Director, Center for Biomedical Image Computing and Analytics, Department of Radiology, University of Pennsylvania, 3600 Market Street, Suite 380, Philadelphia, PA 19104 (christos.davatzikos@uphs.upenn.edu).

†Equally contributing first co-authors.

Background. MRI characteristics of brain gliomas have been used to predict clinical outcome and molecular tumor characteristics. However, previously reported imaging biomarkers have not been sufficiently accurate or reproducible to enter routine clinical practice and often rely on relatively simple MRI measures. The current study leverages advanced image analysis and machine learning algorithms to identify complex and reproducible imaging patterns predictive of overall survival and molecular subtype in glioblastoma (GB).

Methods. One hundred five patients with GB were first used to extract approximately 60 diverse features from preoperative multiparametric MRIs. These imaging features were used by a machine learning algorithm to derive imaging predictors of patient survival and molecular subtype. Cross-validation ensured generalizability of these predictors to new patients. Subsequently, the predictors were evaluated in a prospective cohort of 29 new patients.

Results. Survival curves yielded a hazard ratio of 10.64 for predicted long versus short survivors. The overall, 3-way (long/medium/short survival) accuracy in the prospective cohort approached 80%. Classification of patients into the 4 molecular subtypes of GB achieved 76% accuracy.

Conclusions. By employing machine learning techniques, we were able to demonstrate that imaging patterns are highly predictive of patient survival. Additionally, we found that GB subtypes have distinctive imaging phenotypes. These results reveal that when imaging markers related to infiltration, cell density, microvascularity, and blood–brain barrier compromise are integrated via advanced pattern analysis methods, they form very accurate predictive biomarkers. These predictive markers used solely preoperative images, hence they can significantly augment diagnosis and treatment of GB patients.

Keywords: glioblastoma, imaging, machine learning, predict, survival.

Magnetic resonance imaging (MRI) plays an important role in neuro-oncology for initial diagnosis and assessment of treatment response and is increasingly used as a powerful noninvasive predictive tool. Multiple studies have identified associations between MRI features and survival in patients with high-grade gliomas.[1–8] Common between these diverse analyses is that imaging provides distinct information predictive of survival and outcome independently of pathologic and clinical data.

Similarly, the use of MRI has recently expanded to create noninvasive imaging biomarkers of cellular/molecular characteristics, since imaging phenotypes can be correlated with genomic signatures, and such phenotypes can serve as noninvasive biomarkers of cellular gene expression.[9–11] Thus, MRI can be used to link molecular and imaging diagnostics[12] as a means to further refine and potentially provide for a noninvasive method of diagnosis and prognostication, and for elucidating specific therapeutic targets.[13]

Received 16 March 2015; accepted 12 June 2015

© *The Author(s) 2015. Published by Oxford University Press on behalf of the Society for Neuro-Oncology. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.*

Downloaded from http://neuro-oncology.oxfordjournals.org/ at University of California, Los Angeles on December 9, 2015

Despite the various correlations between imaging features, genomic expression, and survival reported in the literature, no single analysis has been substantive enough to enter clinical practice. Inevitably, considerable difficulty arises when determining which of these imaging biomarkers should be used and, more importantly, which ones will provide the clinician with the most accurate and reproducible result. Three fundamental components are missing from previous work, which limits translation to clinical practice. First, predictive models have been typically designed and tested on the same cohort of patients. These models have not been sufficiently cross-validated, thus decreasing their generalizability (see Supplementary material). Second, only a handful of predefined imaging features are often used, as opposed to synergistically combining all available imaging information to build the most powerful and robust predictive model. Finally, most studies to date have not leveraged the power of modern machine learning technology.[2–4]

To overcome shortcomings of prior studies, we integrate and analyze the entirety of the imaging data, leveraging the power of machine learning and pattern analysis methods. These techniques allow us to extract subtle but distinctive imaging phenotypes, which are not discernible when looking at one or a few imaging variables at a time. Specifically, we hypothesize that overall survival and molecular subtype have an associated imaging phenotype. We test our predictive models retrospectively, using cross-validation, as well as prospectively on new patients, aiming to noninvasively predict a patient's survival and tumor molecular characteristics. Our ultimate goal is to develop robust, noninvasive methodologies to predict clinical and genomic variables of high-grade gliomas from imaging data. In this study, we demonstrate the utility of such an approach in improving the prognostication and molecular subtype identification for patients with glioblastoma (GB).

Materials and Methods

Study Setting and Data Source

We analyzed a retrospective ($N = 105$) followed by a prospective ($N = 29$) cohort of patients with newly diagnosed GB at the University of Pennsylvania from 2006 to 2013. Study inclusion criteria included age over 18 years, histopathological tissue diagnosis of GB (World Health Organization grade IV), and preoperative MRI at time of diagnosis that included structural, diffusion, and perfusion scans. For the molecular subtype study, sufficient tumor tissue collected at time of surgery was required. All patients were treated according to standard of care, which included maximal safe resection, radiotherapy, and concomitant and adjuvant chemotherapy with temozolomide.[14] All patients were followed from time of surgery (typically 2–3 days after baseline MRI) until death to determine overall survival. All causes of death were tumor related. Clinical variables, including age and gender, were extracted from the electronic medical record. Patients were excluded if they had undergone prior GB resection. A separate validation cohort of 29 patients meeting the above inclusion criteria was studied prospectively. All studies were approved by the institutional review board.

Preoperative MRIs were acquired using a 3 Tesla scanner (Magnetom Tim Trio). Obtained for all patients prior to surgery were: T_1-weighted = matrix $192 \times 256 \times 192$; resolution $0.98 \times 0.98 \times 1.00$ mm^3; repetition time (TR): 1760 ms; echo time (TE): 3.1 ms; T_1 with gadolinium (Gd): matrix $192 \times 256 \times 192$; resolution $0.98 \times 0.98 \times 1.00$; TR: 1760 ms; TE: 3.1 ms. T_2-weighted = matrix $208 \times 256 \times 64$; resolution $0.94 \times 0.94 \times 3.00$; TR: 4680 ms; TE: 85 ms. T_2 fluid-attenuated inversion recovery (FLAIR) = matrix $192 \times 256 \times 60$; resolution $0.94 \times 0.94 \times 3.00$; TR: 9420 ms; TE: 141 ms. Diffusion tensor imaging (DTI): matrix $128 \times 128 \times 40$; resolution $1.72 \times 1.72 \times 3.00$; 30 gradient directions. Dynamic susceptibility contrast-enhanced (DSC)-MRI, gradient echo type echo planar imaging (GRE EPI) = field of view (FOV) 22 cm $128 \times 128 \times 20$; resolution $1.72 \times 1.72 \times 3$ mm^3; TR: 2000 ms; TE: 45 ms. An initial loading dose of one-quarter of the total contrast dose was administered first to help minimize errors due to potential contrast leakage out of intravascular space, and DSC-MRI data were acquired during a second bolus of the remaining contrast dose after a 5-min delay for a total of 0.3 mL/kg or 1.5 times single dose MultiHance (gadobenate dimeglumine). For post-processing, blood volume maps were created on a Leonardo workstation (Siemens) using the Neuro Perfusion Evaluation task card as per clinical routine. All MRIs of each patient were coregistered, smoothed, corrected for magnetic field inhomogeneities, and skull stripped, as previously described.[15–17]

Tissue Subtyping

After a formal neuropathologist's (M.M-L.) GB diagnosis, the tissue sample underwent subtyping into one of 4 molecular classes of GB (classical, mesenchymal, proneural, and neural), as described by Verhaak et al.[18] We used an isoform-based classifier initially constructed using exon array data from GB samples in The Cancer Genome Atlas; it was then translated into a clinically applicable platform in which expression of desired transcripts was measured using reverse transcriptase–quantitative polymerase chain reaction (RT-qPCR).[19] RNA was isolated from the tissue samples using Tri Reagent (Sigma). A high-capacity complementary DNA reverse transcriptase kit (Applied Biosystems) was used to reverse-transcribe the RNA, and qPCR was then performed to designate the subtype. The assay was based on the expression of 121 transcripts with four housekeeping genes as controls.

MRI Features

The computer-based glioma image segmentation and registration (GLISTR) segmentation algorithm[20] was used to segment enhancing tumor (ET), non-ET (nonenhancing core of tumor), edema, and ventricles, as well as to estimate parameters of a biophysical glioma growth model[20] and to coregister patient data with a standardized atlas coordinate system for tumor spatial location estimation. Following image segmentation, structural (T_1, T_1-Gd, T_2, T_2-FLAIR, DTI) and physiological (DSC-MRI) imaging sequences were analyzed to select relevant imaging features used in creating a predictive model. Derived from DTI were fractional anisotropy (FA), radial diffusivity (RAD), axial diffusivity (AX), and trace (TR), and derived from DSC-MRI were relative cerebral blood volume (rCBV), peak height (PH), and percentage signal recovery (PSR).

For each GB patient, we extracted multiple features from MRIs in order to capture various phenotypic characteristics of

Downloaded from http://neuro-oncology.oxfordjournals.org/ at University of California, Los Angeles on December 9, 2015

ET, non-ET, and edema. Selected features reflect intensity distributions of various MRI sequences, size and location of the tumor and regions of edema, and parameters extracted from a patient-specific biophysical tumor growth model (which was individually fitted to each patient's images). All features were integrated via a machine learning model that determined which feature combinations were most predictive of survival and molecular subtype.

The initial features included: size and normalized size of enhancing tumor, nonenhancing tumor, edema, and their combinations (14 features); GLISTR outputs: diffusion time, proliferation coefficient, mass-effect parameters, diffusion coefficients of white and gray matter (5 features); distance of tumor to ventricles and distance of edema to ventricles (2 features); mean and standard deviation of intensities of T_1, T_2, T_1-Gd, T_2-FLAIR, rCBV, PH, PSR, FA, RAD, AX, and TR in ET, non-ET, and edema (66 features); frequency of intensities of T_1, T_2, T_1-Gd, FLAIR, rCBV, PH, PSR, FA, RAD, AX, and TR in each distribution bin in ET, non-ET, and edema (330 features); location of the tumor in the brain (9 features); age; and gender.

The distribution of intensities were initially divided into 10 bins and then modified to a lower number of bins based on each modality in each region. Therefore, 330 features of intensity distribution were distilled to 120 features.

Feature Selection

All features were initially tested for their individual predictive value. Feature selection sequentially selected features that best predicted survival until there was no improvement. After adding each candidate feature, 10-fold cross-validation was performed by repeatedly calculating accuracy with different training subsets of patients. About 60 of the most helpful features were finally retained for making each predictive model.

Machine Learning Algorithm

A multidimensional pattern classification method known as Support Vector Machines (SVM)[21] was used in the survival study to construct 2 classifiers. One SVM was used to distinguish between patients surviving less/more than 6 months (short survival) and another SVM was used to distinguish between patients who survived less/more than 18 months (long survival). These 2 indices were then combined to form a single survival prediction index (SPI); higher SPI values indicated longer predicted survival. For predicting subtype, a separate classifier was trained to discriminate among each of the 4 individual GB molecular subtypes and the remaining 3, and voting determined the final 4-way classification.

Evaluating Histograms of Features

In order to obtain a deeper biological understanding of the information used by the machine learning algorithms to provide predictions, histograms of the most predictive features were generated by dividing the patient cohort into 2 groups according to median survival and displaying long/short survivor features separately in blue/red (Fig. 1).

We used 2 types of features. The first one had a single value for each patient: volumes of tumor and ET/non-ET/edema, diffusion time based on a biophysical model of tumor growth,[20] distance of tumor from the ventricle, and age. The vertical axis in these histograms showed the number of patients. Tumor size referred to the size of ET plus non-ET regions.

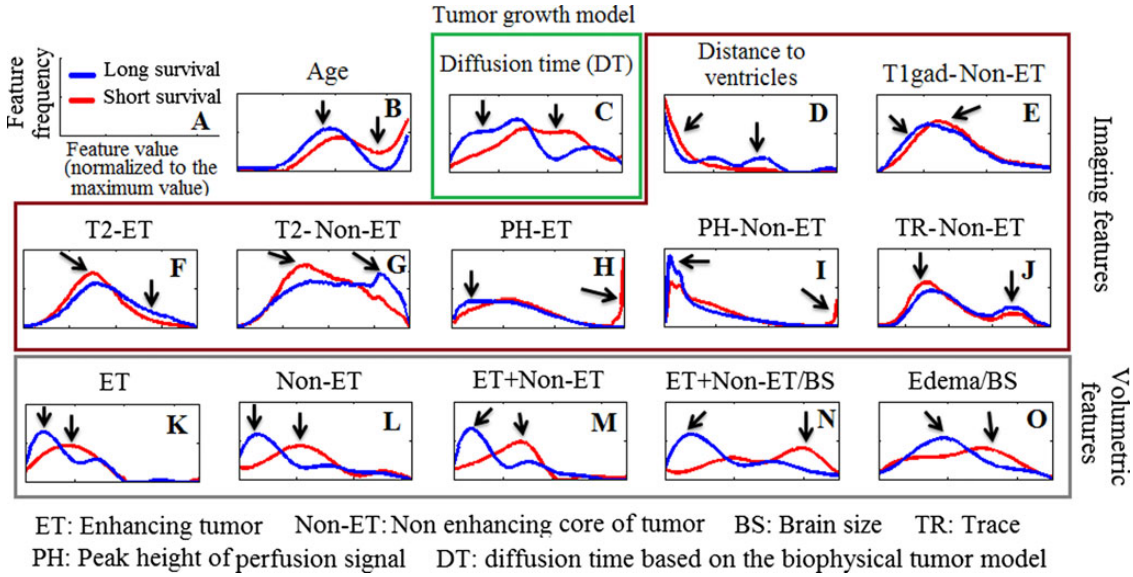

Fig. 1. Histograms display the distribution of features most predictive of survival. The figures show the relationship between the value of each feature (*x*-axis) and frequency of each feature value (*y*-axis) in short (red) and long (blue) survival (median split of patient survival). The black arrows point to marked differences between short and long survivors. (A) Legends; (B) age; (C) diffusion time based on a biophysical model of tumor growth and diffusion; (D) distance of tumor from the ventricles; (E) T_1-Gd intensity in non-ET; (F) T_2 intensity in ET; (G) T_2 intensity in non-ET; (H) peak height (PH) of perfusion signal in ET; (I) PH in non-ET; (J) trace (TR) in non-ET; (K) volume of ET; (L) volume of non-ET; (M) tumor volume (ET plus non-ET); (N) normalized (for total brain size) tumor volume; (O) normalized volume of edema.

Downloaded from http://neuro-oncology.oxfordjournals.org/ at University of California, Los Angeles on December 9, 2015

Normalized size referred to measured size relative to total brain size. The second type of feature was voxel based and displayed intensity distributions of various MRI sequences. Histograms of these features were created with the value of the feature on the *x*-axis and its frequency on the *y*-axis. These histograms were normalized for total number of voxels, in order to account for size effects (which were measured separately) and to be able to appreciate the shape of these histograms, which to a large extent capture the heterogeneity of imaging features within edema, non-ET, and ET.

Retrospective Cohort

Tenfold cross-validation was used to determine the generalization accuracy of the predictive models. In this approach, the model was generated using data for 90% of the subjects. The predictions were then tested on the remaining 10% of subjects. This process was reiterated 10 times, each time leaving out a different 10%.

Prospective (Replication) Cohort

We tested our predictive model prospectively in a cohort of GB patients separate from the cohort that was used to develop it. All subjects met inclusion criteria identical to those of the retrospective cohort study. Such a replication cohort study was performed in order to provide realistic estimates of how well the predictive models were likely to generalize to new patients.

Statistical Analysis

Kaplan–Meier estimates were used to evaluate survival predictions. The difference among 3 survival groups predicted by our model was assessed by measuring the number of patients surviving over a period of time (Fig. 2). The horizontal axis shows the time and the vertical axis shows the probability of survival. Any point on the survival curve shows the probability that a patient in each group would remain alive at that time. Survival curves were compared statistically by a Cox proportional hazards model[22] for statistical significance. All statistical analyses were performed using SPSS (IBM). Hazard ratios and Kaplan–Meier curves were computed for survival analysis, with the level of statistical significance for a 2-sided comparison set at 5% ($P < .05$).

Accuracy of predictions was obtained by dividing the sum of true positives and true negatives by the total number of patients. The overall accuracies for the 3 groups in the survival study and for the 4 groups in the molecular subtype study were calculated based on classification scores obtained via voting (eg, for the molecular subtype, we had 4 classifications for each person, and the most likely defined the classification). Receiver operating characteristic (ROC) curves were calculated as usual, by varying the classification threshold.

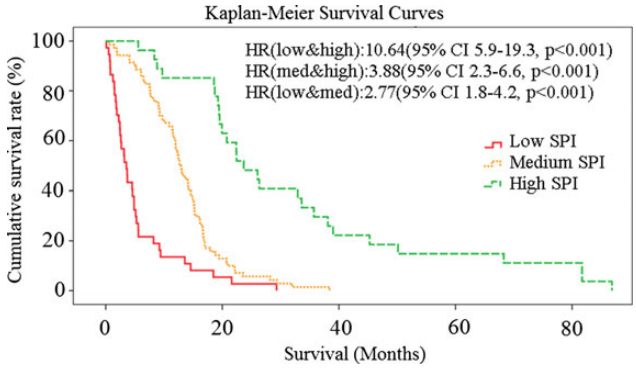

Fig. 2. Kaplan–Meier survival curves. Actual survival on *x*-axis is compared among each of the 3 survival groups based on predictions generated by the survival prediction index (SPI). HR, hazard ratio; med, medium.

Results

Imaging Features

A linear SVM[21] was used to create a SPI that determined whether patients were more likely to survive less than 6 months (short survival), between 6 and 18 months (medium survival), or greater than 18 months (long survival). For molecular subtyping, the machine learning algorithm was trained to recognize one of the 4 GB subtypes (neural, proneural, classical, and mesenchymal).[18,19]

Histograms for the most predictive imaging features regarding survival are in Fig. 1. Histograms have been normalized so that the area under the curve is fixed, to allow for extraction of histogram shape attributes. The features were categorized into different groups: age (Fig. 1B); diffusion time based on a biophysical model of tumor growth fitted to each individual patient (Fig. 1C); imaging features (Fig. 1D–J) including distance of tumor from the ventricles, T_1-Gd intensity in non-ET, T_2 intensity in ET, T_2 intensity in non-ET, PH of perfusion signal in ET, PH in non-ET, TR in non-ET; and volumetric features (Fig. 1K–O) including volume of ET, volume of non-ET, volume of tumor (ET plus non-ET), normalized (for total brain size) tumor volume, and normalized volume of edema. Relatively higher age, volume of ET, non-ET, tumor (ET and non-ET), volume of edema, and relatively smaller distance between tumor and ventricles were predictive of short survival. Patients with a relatively larger number of voxels of high T_1-Gd intensity, low T_2 intensity, high PH, and low TR had relatively shorter survival.

Patient Survival Predictions

Classification rates obtained using the baseline SPI are shown in Table 1. The cross-validated 2-class accuracy of the 6-month survival SVM model was 87.62% in the retrospective study and 82.76% in the prospective study. The cross-validated 2-class accuracy of the 18-month survival SVM model was 88.57% in the retrospective study and 83.33% in the prospective study. Overall 3-way classification into short/medium/long survivors was approximately 77% and 79% in the cross-validated retrospective and prospective cohorts, respectively. Addition of tumor molecular subtype—estimated from RT-qPCR data according to the Tissue Subtyping section of this paper—to the feature set did not improve survival classification accuracy.

Survival Curves

Kaplan–Meier survival curves for each of the 3 groups based on the SPI are shown in Fig. 2, with the *x*-axis representing

Downloaded from http://neuro-oncology.oxfordjournals.org/ at University of California, Los Angeles on December 9, 2015

Table 1. Accuracy of molecular subtype and survival prediction; the ROC curve is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied

	Study (*n*)	Accuracy (%)	AUC
Survival			
<6 mo	Retrospective (27)	87.62	0.87
	Prospective (11)	82.76	0.85
>18 mo	Retrospective (29)	88.57	0.91
	Prospective (7)	83.33	0.84
6–18 mo	Retrospective (49)	79.05	N/A
	Prospective (11)	87.5	N/A
Overall	Retrospective (105)	77.14	N/A
	Prospective (29)	79.17	N/A
Tumor subtype	Proneural (22)	85.86	0.87
	Neural (29)	87.88	0.92
	Mesenchymal (28)	83.84	0.89
	Classical (20)	84.85	0.75
	Overall (99)	75.76	N/A

Therefore, AUC only can be calculated for the classification of 2 groups.

actual survival in months (both cross-validated retrospective and prospective data are included). Actual survival was highest among patients predicted by the model to be long survivors (highest SPI) and lowest among those predicted to be short survivors. Intermediate survivors consisted of subjects predicted to survive more than 6 months and less than 18 months. The hazard ratio between low and high SPI was 10.64 (95% CI = 5.87–19.28, $P < .001$; the hazard ratio between intermediate and high SPI was 3.88 (95% CI = 2.27–6.64, $P < .001$); and the hazard ratio between intermediate and short SPI was 2.77 (95% CI = 1.83–4.18, $P < .001$. Figure 3 shows box plots of SPI versus actual survival (higher SPI predicts longer survival).

Cutoff Values of Important Survival Features

Although to achieve predictions reported herein one would need all the features and our predictive model, we now attempt to provide simplified interpretations and thresholds that can be used clinically with relative ease. Specifically, patients in the lower 50th percentile for survival (median survival, 12.9 mo) tended to have the following characteristics: (i) age higher than 64.6 (area under the curve [AUC], 0.7); (ii) tumor volume bigger than 1.6% of cerebral hemispheric volume (AUC, 0.65); (iii) diffusion time longer than 100 units (more simply, tumors with diffusive appearance have poorer prognosis); (iv) the lower half of the T_2-intensity histogram of ET voxels containing more than 42.7% of the voxels (AUC, 0.62); (v) after binning the non-ET voxels' T_2 intensities into 3 bins, the higher bin containing less than 38.2% of the non-ET voxels (AUC, 0.56); (vi) having more than 1% of ET voxels in the brightest tenth bin of peak height of the perfusion signal (put simply, spots of highest perfusion peak height in ET—and to a lesser extent in non-ET—is a nearly certain indication of short survival).

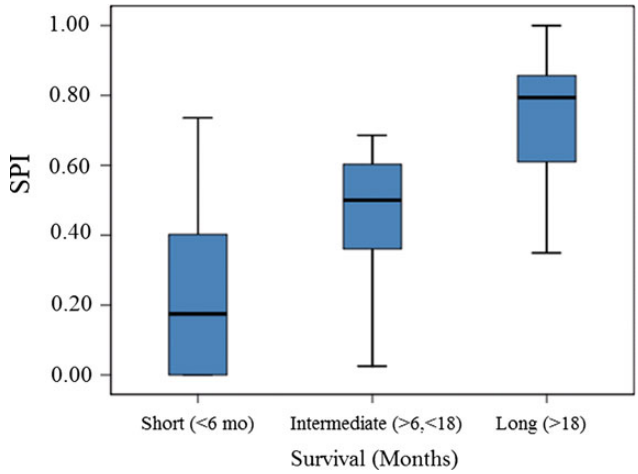

Fig. 3. SPI versus actual survival. Solid transverse black line is the median, edges of boxes are first and third quartiles, and error bars are ±1.5 interquartile range.

Predicting Tumor Molecular Subtype

A subgroup of 99 patients who had sufficient tumor tissue collected at time of surgery was included in the molecular subtype study and underwent subtyping using an isoform-level assay classifier.[19] The numbers of classical, mesenchymal, proneural, and neural subtypes were 20, 28, 22, and 29, respectively. Using feature selection, a unique predictive algorithmic classifier was then built for each tumor subtype. Table 1 shows the accuracy of each subtype model, as well as the overall accuracy in the 4-way classification. Figure 4 demonstrates the ROC curves for molecular subtype classifiers. Image-derived features most predictive of molecular subtype included histogram of T_2-FLAIR intensity in ET, size of ET, and PH on perfusion signal in edema for the classical subtype; mean T_1 signal in ET and histogram of T_2-FLAIR intensity in edema for the mesenchymal subtype; T_2-intensity histogram in edema and tumor location for the neural subtype; and histogram of T_2-FLAIR intensity in ET and mean T_1 in ET for the proneural subtype.

Discussion

By integrating a rich set of imaging features using machine learning algorithms, we achieved accurate and reproducible prediction of patient survival using preoperative MRI. Importantly, these survival estimates were obtained using cross-validation retrospectively and were subsequently prospectively tested on an entirely new validation cohort of patients. The use of such cross-validation techniques bolsters generalizability and ensures that this model and finding will replicate in future studies. Using a similar approach, we predicted individual GB molecular subtypes with an overall accuracy of 76%. These findings support our initial hypothesis that genomic events, in this case the molecular subtypes of GB, have unique imaging phenotypes. Notably, these predictions could not be achieved by any individual MRI feature and became possible only after integrating diverse imaging information using machine learning methods.

Downloaded from http://neuro-oncology.oxfordjournals.org/ at University of California, Los Angeles on December 9, 2015

Predicting Survival

Our image-based predictive model of overall survival improves upon approaches used in prior prognostic imaging studies in GB. Traditionally, clinical variables such as age, Karnofsky performance score, and extent of resection have been used to predict overall survival.[23] The inclusion of imaging features, however, improves upon survival prediction compared with the use of clinical features alone.[5] For example, image analysis has identified variables such as the ratio of T_2-FLAIR signal to tumor volume,[4] proportion of ET,[7,24] spatial variation in enhancement,[4] and volume of ET crossing the corpus callosum[25] to strongly correlate with survival. These previous attempts at correlating imaging phenotypes with survival have focused on one or a handful of imaging features, as opposed to integrating multiple variables. This has led to significant variation in the literature defining the imaging feature(s) most predictive of overall survival. Furthermore, the majority of these studies did not test the reproducibility of the fitted regression models in a prospective manner or via cross-validation.

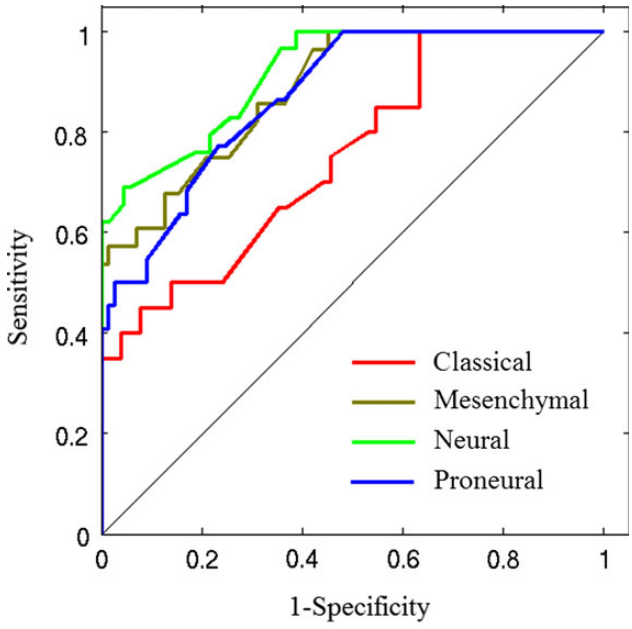

Fig. 4. ROC curves for each glioblastoma molecular subtype prediction are compared with chance (the diagonal line). AUC is shown in Table 1.

Automated segmentation and machine learning techniques[4,24] have facilitated the assessment of multiple imaging and clinical features simultaneously. These techniques are particularly important when analyzing large datasets, as they can identify complex patterns in data that simple regression models obscure. We believe that our study is the most comprehensive to date; we used over 300 clinical and derived imaging features to construct our predictive models. This type of comprehensive approach is critical, as no single variable is predictive of survival in a consistent and reproducible manner, especially when tested prospectively. As is evidenced by the histograms for individual features (Fig. 1), small differences frequently exist between short and long survivors, and it is only through the integration of multiple features that highly distinctive patterns predictive of survival emerge.

In addition to standard anatomic imaging, we leveraged physiological imaging modalities (perfusion MRI and DTI) to gain insight into tumor biology and further augment our predictive models. Features derived from DSC-MRI and dynamic contrast-enhanced (DCE) MRI have been previously used to develop predictive models for survival in patients with high-grade gliomas.[2,6,26,27] These studies identified biomarkers for angiogenesis (rCBV) and cellularity (apparent diffusion coefficient [ADC]) as being important for differentiating short and long survivors.[28–30] Our study goes beyond these relatively focused studies by using an extensive set of features derived from perfusion and diffusion sequences and integrating them with features derived from structural MRI.

The imaging features identified by our analyses to be most predictive of survival relate to the underlying pathophysiology of GB (Table 2), characterized by angiogenesis, breakdown of the blood–brain barrier, and peritumoral infiltration. These cellular events have been shown to have specific imaging phenotypes. For example, Gd enhancement reflects blood–brain barrier breakdown,[31,32] DSC-MRI is able to estimate tissue microvascularity,[33,34] and diffusion TR relates to tissue

Table 2. Imaging features and biological interpretation

Predictive Feature Class	Finding and Interpretation
Age	Higher age → shorter survival
Volumetrics: ET, non-ET, normalized edema	Larger and more infiltrating tumor → shorter survival
Water concentration/cell density (T_2 in ET and non-ET, diffusion TR in non-ET)	Higher water concentration and lower cell density → longer survival
Microvascularity (PH in ET, PH in non-ET, T_1-Gd in non-ET)	Increased and compromised tumor microvascularity → shorter survival
Biophysical model	Higher diffusion time → shorter survival (suggestive of diffusive tumors having poorer outcome)
Spatial location (distance from the ventricles)	Proximity to the ventricles leads to relatively shorter survival. Possibly reflects poorer outcome of patients with tumors extending into inoperable regions (higher extent of critical region involvement), or potentially related to access to subventricular stem-cell-generating regions leading to poorer outcome.

More detailed information is shown in Fig. 1.

Downloaded from http://neuro-oncology.oxfordjournals.org/ at University of California, Los Angeles on December 9, 2015

cellularity.[35] In our survival model, patients who were predicted to survive longer had relatively lower volumes of ET, relatively lower tissue volumes with increased PH (PH relates to increased and compromised microvascularity), and lower cellular density/higher fluid concentration (eg, higher TR, T_2 signal). Thus, consistent with the pathophysiology of GB, our image-based model predicted that short survivors would have imaging features consistent with increased angiogenesis, breakdown of the blood–brain barrier, increased peritumoral infiltration and cell density, and decreased water content. Although each of these imaging features provides complementary information regarding the biological properties of the underlying tissue, the combination of multiple features in an integrative model yields stronger predictions than a model based on any single feature alone, which once again reflects the main strength of our study.

Additional imaging features found to be predictive of short survival in our model relate to increased tumor burden, growth, and location. As in prior studies,[4,5,7] normalized edema volume, ET volume, non-ET volume, tumor volume, and normalized tumor volume were larger in short survivors compared with long survivors (Fig. 1). Additionally, diffusion time, a feature derived from the biophysical growth model felt to represent invasiveness over time,[20,36,37] was higher in short survivors (Fig. 1). Finally, distance of the tumor from the ventricle was shorter and age was higher in patients with shorter survival (Fig. 1). The predictive value of distance from ventricles could reflect either the limited ability to surgically resect deep tumors or the tumor's access to the stem-cell-generating subventricular zone leading to poorer outcome (see Supplementary material).

Predicting Molecular Subtype

In addition to prognostic information, MRI features may also serve as noninvasive biomarkers of underlying molecular events. Several authors have demonstrated that specific GB molecular subtypes have unique imaging features that can be exploited. For instance, mesenchymal tumors have lower non-ET volume and peritumoral edema intensity, the proneural subtype has significantly lower blood–brain barrier breakdown, and the classical subtype strongly correlates with necrosis and edema imaging features related to edge sharpness and intensity.[7,38] In fact, a model to predict the mesenchymal subtype has been previously suggested; however, the authors were unable to find distinguishing imaging features among the non-mesenchymal subtypes.[39] Whereas prior studies have correlated structural imaging features with a single GB molecular subtype, to our knowledge our study is the first to noninvasively classify a GB patient into one of the 4 individual subtypes with relatively high accuracy based solely on routinely obtained clinical imaging data.

Multiple studies focus on genomic predictors of survival in GB patients. In our model, GB molecular subtype was assessed,[18] but it was not found to offer predictive value beyond what was predicted via imaging. Although classification of molecular subtype alone has not been shown to predict survival, the addition of subtype data can augment predictive models. For instance, Jain et al[40] found that rCBV alone, as measured by DSC-MRI, was not predictive of survival. However, when the molecular subclass was added to the model, maximum rCBV became a significant survival predictor. Thus, subtype classification may provide additional information, especially as an adjunct to hemodynamic imaging biomarkers. Unlike the study by Jain et al in which molecular subtype was obtained through tissue analysis, we have shown in the current study that we were able to extract such information through imaging alone.

Limitations

Despite precautions, including use of multiple imaging features, automated region of interest measurements, and cross-validation of all results, our study has limitations. First, we used only the MRI at time of diagnosis in creating our predictive model, rather than incorporating postoperative imaging. Other groups have demonstrated, for instance, that peritumoral perfusion parameters on posttreatment imaging, not preoperative imaging, were associated with overall survival,[41] while others have performed serial analyses of MRI.[10] We intentionally chose to evaluate only preoperative imaging in order to avoid any potential confounding factors related to treatment effects that may manifest on postoperative imaging. For example, the presence of blood and the reduction in tumor volume can alter TR measurements and lead to false-positive results,[30] It is also important to note that all patients in our study received the same clinical treatment: surgical debulking followed by chemo and radiotherapy. Another limitation of our study is that we used data from a single institution; data from other institutions would be beneficial to further validate our hypothesis and the external validity of our predictive models.

Significance

Using MRIs at time of diagnosis, we applied machine learning algorithms to predict noninvasively, and with high accuracy, clinical and genomic variables of patients with brain cancer. Our findings therefore provide a robust mechanism for predicting survival among GB patients and for estimating molecular subtype. By utilizing standard imaging sequences that are commonly employed in clinical practice and by testing our models prospectively, our findings can be readily translated to the clinic. These predictive models may be used by clinicians to offer prognostic survival data to patients with newly diagnosed disease as well as guide the aggressiveness of treatment. Additionally, the molecular subtype information can guide the potential enrollment of patients into experimental treatments upfront. Furthermore, GB tumors exhibit regional differences in genomic alterations, enhancement, cell density, and necrosis,[42] and imaging provides a means of capturing such spatial heterogeneity, especially when pathology sampling is suboptimal or not possible due to tumor location. Thus, imaging enables the detection of phenotypic signatures that may not be captured by classical histopathology or other individual biomarkers.

Supplementary Material

Supplementary material is available at *Neuro-Oncology Journal* online (http://neuro-oncology.oxfordjournals.org/).

Downloaded from http://neuro-oncology.oxfordjournals.org/ at University of California, Los Angeles on December 9, 2015

Funding

This work was supported in part by NIH grant R01NS 042645.

Conflict of interest statement. Nothing to disclose.

References

1. Akbari H, Macyszyn L, Da X, et al. Pattern analysis of dynamic susceptibility contrast-enhanced MR imaging demonstrates peritumoral tissue heterogeneity. *Radiology*. 2014;273(2):502–510.
2. Law M, Young RJ, Babb JS, et al. Gliomas: predicting time to progression or survival with cerebral blood volume measurements at dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging 1. *Radiology*. 2008;247(2):490–498.
3. Pope WB, Qiao XJ, Kim HJ, et al. Apparent diffusion coefficient histogram analysis stratifies progression-free and overall survival in patients with recurrent GBM treated with bevacizumab: a multi-center study. *J Neurooncol*. 2012;108(3):491–498.
4. Zhang Z, Jiang H, Chen X, et al. Identifying the survival subtypes of glioblastoma by quantitative volumetric analysis of MRI. *J Neurooncol*. 2014;119(1):207–214.
5. Mazurowski MA, Desjardins A, Malof JM. Imaging descriptors improve the predictive power of survival models for glioblastoma patients. *Neuro Oncol*. 2013;15(10):1389–1394.
6. Bonekamp D, Deike K, Wiestler B, et al. Association of overall survival in patients with newly diagnosed glioblastoma with contrast-enhanced perfusion MRI: comparison of intraindividually matched T1-and T2*-based bolus techniques. *J Magn Reson Imaging*. 2014;42(1):87–96.
7. Gutman DA, Cooper LA, Hwang SN, et al. MR imaging predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set. *Radiology*. 2013;267(2):560–569.
8. Li W-b, Kai T, Chen Q, et al. MRI manifestions correlate with survival of glioblastoma multiforme patients. *Cancer Biol Med*. 2012;9(2):120–123.
9. Diehn M, Nardini C, Wang DS, et al. Identification of noninvasive imaging surrogates for brain tumor gene-expression modules. *Proc Natl Acad Sci U S A*. 2008;105(13):5213–5218.
10. Li Y, Lupo JM, Polley M-Y, et al. Serial analysis of imaging parameters in patients with newly diagnosed glioblastoma multiforme. *Neuro Oncol*. 2011;13(5):546–557.
11. Tykocinski ES, Grant RA, Kapoor GS, et al. Use of magnetic perfusion-weighted imaging to determine epidermal growth factor receptor variant III expression in glioblastoma. *Neuro Oncol*. 2012;14(5):613–623.
12. Rutman AM, Kuo MD. Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. *Eur J Radiol*. 2009;70(2):232–241.
13. Zinn PO, Majadan B, Sathyan P, et al. Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme. *PloS One*. 2011;6(10):e25451.
14. Stupp R, Hegi ME, Mason WP, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol*. 2009;10(5):459–466.
15. Jenkinson M, Beckmann CF, Behrens TE, et al. FSL. *Neuroimage*. 2012;62(2):782–790.
16. Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. *Med Image Anal*. 2001;5(2):143–156.
17. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002;17(3):143–155.
18. Verhaak RG, Hoadley KA, Purdom E, et al. Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*. 2010;17(1):98–110.
19. Pal S, Bi Y, Macyszyn L, et al. Isoform-level gene signature improves prognostic stratification and accurately classifies glioblastoma subtypes. *Nucleic Acids Res*. 2014;42(8):e64.
20. Gooya A, Biros G, Davatzikos C. Deformable registration of glioma images using EM algorithm and diffusion reaction modeling. *IEEE Trans Med Imaging*. 2011;30(2):375–390.
21. Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*. 2011;2(3), Article 27.
22. Cox DR. Regression models and life-tables. *J R Stat Soc B* 1972;34(2):187–220.
23. Lacroix M, Abi-Said D, Fourney DR, et al. A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival. *J Neurosurg*. 2001;95(2):190–198.
24. Mazurowski MA, Zhang J, Peters KB, et al. Computer-extracted MR imaging features are associated with survival in glioblastoma patients. *J Neurooncol*. 2014;120(3):483–488.
25. Ramakrishna R, Barber J, Kennedy G, et al. Imaging features of invasion and preoperative and postoperative tumor burden in previously untreated glioblastoma: correlation with survival. *Surg Neurol Int*. 2010;1:40.
26. Chaskis C, Stadnik T, Michotte A, et al. Prognostic value of perfusion-weighted imaging in brain glioma: a prospective study. *Acta Neurochir (Wien)*. 2006;148(3):277–285.
27. Hirai T, Murakami R, Nakamura H, et al. Prognostic value of perfusion MR imaging of high-grade astrocytomas: long-term follow-up study. *Am J Neuroradiol*. 2008;29(8):1505–1510.
28. Saksena S, Jain R, Narang J, et al. Predicting survival in glioblastomas using diffusion tensor imaging metrics. *J Magn Reson Imaging*. 2010;32(4):788–795.
29. Crawford FW, Khayal IS, McGue C, et al. Relationship of pre-surgery metabolic and physiological MR imaging parameters to survival for patients with untreated GBM. *J Neurooncol*. 2009;91(3):337–351.
30. Nakamura H, Murakami R, Hirai T, et al. Can MRI-derived factors predict the survival in glioblastoma patients treated with postoperative chemoradiation therapy? *Acta radiol*. 2013;54(2):214–220.
31. Kalpathy-Cramer J, Gerstner ER, Emblem KE, et al. Advanced magnetic resonance imaging of the physical processes in human glioblastoma. *Cancer Res*. 2014;74(17):4622–4637.
32. Kelly PJ, Daumas-Duport C, Kispert DB, et al. Imaging-based stereotaxic serial biopsies in untreated intracranial glial neoplasms. *J Neurosurg*. 1987;66(6):865–874.
33. Hu L, Baxter L, Smith K, et al. Relative cerebral blood volume values to differentiate high-grade glioma recurrence from posttreatment

Downloaded from http://neuro-oncology.oxfordjournals.org/ at University of California, Los Angeles on December 9, 2015

radiation effect: direct correlation between image-guided tissue histopathology and localized dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging measurements. *Am J Neuroradiol*. 2009;30(3):552–558.

34. Pathak AP, Schmainda KM, Ward BD, et al. MR-derived cerebral blood volume maps: issues regarding histological validation and assessment of tumor angiogenesis. *Magn Reson Med*. 2001; 46(4):735–747.
35. Beppu T, Inoue T, Shibata Y, et al. Fractional anisotropy value by diffusion tensor magnetic resonance imaging as a predictor of cell density and proliferation activity of glioblastomas. *Surg Neurol*. 2005;63(1):56–61.
36. Swanson KR, Alvord E, Murray J. Virtual brain tumours (gliomas) enhance the reality of medical imaging and highlight inadequacies of current therapy. *Br J Cancer*. 2002;86(1):14–18.
37. Swanson KR, Bridge C, Murray J, et al. Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. *J Neurol Sci*. 2003;216(1):1–10.
38. Gevaert O, Mitchell LA, Achrol AS, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. *Radiology*. 2014;273(1):168–174.
39. Naeini KM, Pope WB, Cloughesy TF, et al. Identifying the mesenchymal molecular subtype of glioblastoma using quantitative volumetric analysis of anatomic magnetic resonance images. *Neuro Oncol*. 2013;15(5):626–634.
40. Jain R, Poisson L, Narang J, et al. Genomic mapping and survival prediction in glioblastoma: molecular subclassification strengthened by hemodynamic imaging biomarkers. *Radiology*. 2013;267(1): 212–220.
41. Bag AK, Cezayirli PC, Davenport JJ, et al. Survival analysis in patients with newly diagnosed primary glioblastoma multiforme using pre-and post-treatment peritumoral perfusion imaging parameters. *J Neurooncol*. 2014;120(2):361–370.
42. Zhou M, Hall L, Goldgof D, et al. Radiologically defined ecological dynamics and clinical outcomes in glioblastoma multiforme: preliminary results. *Transl Oncol*. 2014;7(1):5–13.

Downloaded from http://neuro-oncology.oxfordjournals.org/ at University of California, Los Angeles on December 9, 2015