MULTIVARIATE PATTERN ANALYSIS OF IN VIVO MR IMAGING PARAMETERS FOR DETECTING TRANSFORMATIONS TO A HIGHER GRADE IN PATIENTS WITH RECURRENT LOW GRADE GLIOMAS

Alexandra Constantin[1], Llewellyn Jalbert[1], Adam Elkhaled[2], Rupa Parvataneni[2], Annette Molinaro[2], Joanna Phillips[2], Soonmee Cha[2], Susan M. Chang[2], and Sarah J. Nelson[2]

[1]University of California, Berkeley, Berkeley, CA, United States, [2]University of California, San Francisco

Introduction: The objective of this study was to build a multi-parametric diagnostic model that can predict whether gliomas originally diagnosed a being low grade have undergone malignant transformations to a higher grade, based on *in vivo* magnetic resonance (MR) imaging parameters. This is an important clinical application because it will determine the therapeutic options offered to patients. In some cases this involves surgical resection to provide histological confirmation of tumor grade but in other cases the preferred options are radiation, temozolomide or standard chemotherapy. Having an objective, non-invasive model to predict tumor grade would allow patients to make an informed decision about the risks and benefits of different approaches and could have a major impact upon quality of life and survival.

Data Acquisition: The study population comprised of 61 patients with an original diagnosis of low grade glioma (LGG), who were scheduled for surgical resection due to suspected recurrence. MR examinations were performed in order to provide data for planning image-guided surgery. The protocol included a fluid attenuated inversion recovery (FLAIR) sequence, a T2 weighted fast spin-echo (FSE) sequence, and a T1-weighted sequence obtained after the injection of gadolinium contrast agent (T1-gad). These were used to define the spatial extent of the anatomic lesion. Physiological images that were acquired comprised of 6-directional diffusion tensor imaging (DTI) data with b=1000 mm^2, dynamic susceptibility contrast (DSC) images before, during, and after the injection of 0.1 mmol/kg body weight Gd-DTPA contrast agent at a rate of 3mL/sm. Metabolic imaging data were obtained using lactate-edited 3-D magnetic resonance spectroscopic imaging (MRSI) using point-resolved spectroscopy volume-selection with echo planar and phase encoding localization. Image-guided tissue samples were obtained during surgery and were used to confirm the diagnosis. Histological analysis indicated that 37 of the patients had lesions that progressed to a higher grade (31 to grade 3, and 6 to grade 4), and 24 patients had lesions that remained grade 2.

Preprocessing: The MR images were registered to the T1-gad data and resampled to be in the same orientation. Maps of the apparent diffusion coefficient (ADC) and fractional anisotropy (FA) were obtained from the DTI sequence. Maps of the cerebral blood volume (CBV), peak height (PH), recirculation factor (RF), and percent recovery to baseline (RECOV) were estimated from the DSC susceptibility curves. The amplitudes of peaks corresponding to choline (Cho), creatine (Cre), N-acetyl-aspartate (NAA), lactate (Lac), and lipid (Lip) were estimated from the MRSI data. Regions of interest (ROIs) were created manually and included the abnormalities on FLAIR images (T2all), the contrast enhancing lesions on the T1-gad images (CEL) and any hypointense regions of necrosis on the T1-gad images (NEC). The non-enhancing lesion (NEL) was defined as the difference of the T2all, the CEL, and the NEC. A total of 100 imaging parameters were extracted from these data and used in the analysis. These included volumes of the ROIs and volumes of regions with low ADC, high Cho to NAA (CNI), and high CBV. Also included were metabolite heights and areas inside the T2all, as well as median, 25th percentile, 75th percentile, and maximum imaging intensity values.

Analysis: The problem of identifying lesions that progressed to a higher grade was formulated as a supervised learning problem. The input vector was comprised of the 100 imaging parameters, and the associated grade of recurrence (low grade or high grade) was the desired output or class. The goal was to build a supervised learning model to determine the probability that a lesion recurred at a higher grade for any valid input vector, after having seen a number of training examples, by generalizing from the presented training data to unseen situations in a reasonable way. The type of model we used was additive logistic regression, or boosting [1]. This method uses voting to combine the output of univariate regression models, which are built iteratively and in such a way that new models complement the previous ones by becoming experts for instances handled incorrectly by earlier models. Each model's contribution is weighted by its performance. Suppose f_j is the j^{th} regression model in the ensemble and $f_j(a)$ is its prediction for instance a**.** Then, for a two-class problem, the probability estimate that a belongs to the first class is $p(1 \mid a) = 1/(1+\exp(\sum_j f_j(a)))$ [1]. The variables used as input to the multivariate additive logistic regression model were selected using an automatic, wrapper-based feature selection method [1]. The final model was evaluated using leave-one-out (LOO) cross-validation and bias-corrected bootstrapping. These are the standard validation methods for supervised learning algorithms built on small data sets. LOO repeatedly leaves out one data sample for testing, uses the rest of the samples for training the model, and then averages the accuracy results. It provides an accuracy estimate which is unbiased but has a high variance. Bootstrapping uses sampling with replacement to repeatedly split the data into a training set and a test set, and averages the accuracy results in order to obtain an overall accuracy estimate which is pessimistic, but has low variance.

Results: Model 1 was designed to distinguish between tumors that recurred at a higher grade and those that remained grade 2. This model had 93.02% LOO accuracy with a 95% confidence interval (CI) of [87.84%-98.20%]. The bias-corrected bootstrapping accuracy was 83.59% with a 95% CI of [80.19-86.99]. The model assigned a high probability of progression to lesions with (i) the 75th percentile of normalized choline height above 1.82, (ii) the 25th percentile of recovery to baseline below 76, (iii) the 75th percentile of CNI height above 1.25, and (iv) the maximum choline height above 1.63.

Models (2-5) were designed to compare the utility of different imaging methods for identifying malignant transformation, by building supervised learning models using individual sets of variables selected from (a) MRSI, (b) DSC, (c) DTI and, (d) anatomical imaging parameters. The results are summarized in Table 1. Selecting only MRSI or DSC parameters leads to models with lower accuracies than were obtained by selecting parameters from all imaging sequences, but these differences are not statistically significant. Selecting only DTI or anatomical parameters leads to models with significantly lower accuracy than those obtained using MRSI and DSC parameters.

An analysis of the selected features from the models revealed that all of the astrocytomas (AS) and oligoastrocytomas (OA) with a 75th percentile choline area above 1.065, as well as all oligodendrogliomas (OD) with a 75th percentile choline area above 1.82, progressed to a higher grade. All lesions with a maximum choline height lower than 1.63 remained grade 2, regardless of histological subtype. For AS and OA, a 25th percentile recovery lower than 76 indicated that the lesions progressed to a higher grade. For OD, progression to a higher grade was also indicated by 25th percentile recovery lower than 76 in cases of OD that had a 75th percentile choline area below 1.82. Classification errors were most common among the OD subtype.

Table 1. LOO and bootstrapping accuracy of additive logistic regression models for identifying recurrent LGGs that progressed to a higher grade.

Initial Parameters	LOO acc.	LOO 95% CI	Bootstrapping acc.	Bootstrapping 95% CI	Parameters Selected
Model 1: All	93.02	87.84-98.20	83.59	80.19-86.99	RECOV 25%, Cho area 75%, CNI height 75%, Cho max height
Model 2: MRSI	86.05	77.25-94.85	78.52	75.47-81.57	CNI area >2 volume, CNI area median, Lac median, Cho area 75%
Model 3: DSC	84.88	75.99-93.77	78.82	75.64-82.00	RECOV 25%, PH 75%, CBV>3 volume, median PH
Model 4: DTI	75.58	65.57-85.59	68.80	66.79-70.81	FA median, ADC 75%
Model 5: Anatomy	79.07	69.89-88.25	71.97	69.53-74.41	T1-gad 25%, FLAIR 75%, T1-gad 75%

Discussion: Multivariate pattern recognition methods using parameters obtained from *in vivo* DSC and MRSI data can predict which lesions exhibited malignant transformation with good accuracy. Key parameters in this analysis are high levels of choline and CNI, high intensity values in the contrast-enhancing lesion, low recovery to baseline, large lesion volumes, and large volumes of CBV>2. Choline-based parameters are associated with increased tumor cellularity, and a low recovery to baseline is indicative of abnormal vasculature. After additional validation on an independent data set, the multivariate additive logistic regression model that was built using these data could be used in clinical practice to identify patients suspected of recurrence at a higher grade for cases where histological confirmation of the diagnosis is not possible. The model also suggests that tissue samples obtained for diagnostic purposes should be obtained from areas with high choline, CNI and CBV. While individual *in vivo* parameters were influenced by the histological subtype of the lesion, the choice of parameters and cutoffs selected by the multivariate model are predictive of progression across all histological subtypes.

References: [1]Witten and Frank (2005), Data Mining. [2]Nelson (2003), *Mol Can Therapeutics* 2(5): 497-507. [3]Essock-Burns (2011), *Neuro Oncol* 13(1): 119-131.

This study was supported by NIH Grants CA097257 and CA118816.