Outcome Prediction for Patient with High-Grade Gliomas from Brain Functional and Structural Networks

Luyan Liu[1], **Han Zhang**[2], **Islem Rekik**[2], **Xiaobo Chen**[2], **Qian Wang**[1], and **Dinggang Shen**[2]

[1]Med-X Research Institute, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China

[2]Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, USA

Abstract

High-grade glioma (HGG) is a lethal cancer, which is characterized by very poor prognosis. To help optimize treatment strategy, accurate preoperative prediction of HGG patient's outcome (i.e., survival time) is of great clinical value. However, there are huge individual variability of HGG, which produces a large variation in survival time, thus making prognostic prediction more challenging. Previous brain imaging-based outcome prediction studies relied only on the imaging intensity inside or slightly around the tumor, while ignoring any information that is located far away from the lesion (i.e., the "normal appearing" brain tissue). Notably, in addition to altering MR image intensity, we hypothesize that the HGG growth and its mass effect also change both structural (can be modeled by diffusion tensor imaging (DTI)) and functional brain connectivities (estimated by functional magnetic resonance imaging (rs-fMRI)). Therefore, integrating connectomics information in outcome prediction could improve prediction accuracy. To this end, we unprecedentedly devise a machine learning-based HGG prediction framework that can effectively extract valuable features from complex human brain connectome using network analysis tools, followed by a novel multi-stage feature selection strategy to single out good features while reducing feature redundancy. Ultimately, we use support vector machine (SVM) to classify HGG outcome as either *bad* (survival time 650 days) or *good* (survival time >650 days). Our method achieved 75 % prediction accuracy. We also found that functional and structural networks provide complementary information for the outcome prediction, thus leading to increased prediction accuracy compared with the baseline method, which only uses the basic clinical information (63.2 %).

1 Introduction

Gliomas account for around 45 % of primary brain tumors. The prognosis of gliomas depends on multiple factors, such as age, histopathology, tumor size and location, extent of resection, and type of treatment. Most deadly gliomas are classified by World Health Organization (WHO) as Grade III (anaplastic astrocytoma, and anaplastic oligodendroglioma) and Grade IV (glioblastoma multiforme), according to the histopathological subtypes. These are referred to as high-grade gliomas (HGG) with fast

---

Correspondence to: Dinggang Shen.

growing rate and diffusive infiltration. More importantly, HGG are characterized by a very poor prognosis but the outcome (i.e., the overall survival time) differs significantly from case to case. This can be explained by a large variation in tumor characteristics (e.g., location, and cancer cell type). Yet still challenging, pre-operative prediction of HGG outcome is of great importance and is highly desired by clinicians.

Multimodal presurgical brain imaging has been gaining more solid ground in surgical planning. In turn, this produces abundant multimodal neuroimaging information for potential HGG outcome prediction. For instance, in [1], multiple features, reflecting intensity distributions of various magnetic resonance imaging (MRI) sequences, were extracted to predict patient survival time and molecular subtype of glioblastoma. In [2], morphologic features and hemodynamic parameters, along with clinical and genomic biomarkers, were used to predict the outcome of glioblastoma patients. In [3], data mining techniques based on image attributes from MRI produced better HGG outcome prediction performance, than that solely using histopathologic information. Although promising, all these studies shared a first key limitation: they overlooked the relationship between brain connectivity and the outcome. In other words, they mainly relied on extracting information from the tumor region (i.e., tumor and necrotic tissue) or around it (e.g., edema region). This excludes the majority of the "normal appearing" brain tissue — which most likely has been also affected by the tumor. Based on all these information, our hypothesis is rooted in the fact that HGG highly diffuses along white matter fiber tracts, thus altering the brain structural connectivity. Consecutively, altered structural connectivity will lead to functional connectivity. Moreover, the mass effect, edema and abnormal neovascularization may further change brain functional and structural connectivities. Therefore, connectomics data may present useful and complementary information to intensity-based survival time prediction. A second key limitation of previous studies is that none of them compared the prediction performance when using conventional clinical data *versus* when using advanced connectomics data from multimodalities. We aimed to address both of the limitations.

Conventional neuroimaging computing methods, such as graph-theory-based complex network analysis, have demonstrated promising value in disease classification and biomarker detection [4]. However, to our best knowledge, no previous study has utilized brain connectome to predict the treatment outcome for HGG patients. In this work, we hypothesize that gliomas have 'diffusive effects' to both structural and functional connectivities, involving both white matter and grey matter, which could alter the inherent brain connectome and lead to abnormalities in network attributes. Hence, we devise an HGG outcome prediction framework, by integrating, extracting and selecting the best set of advanced brain connectome features.

Specifically, we retrospectively divided the recruited HGG patients into *short* and *long* survival time groups based on the follow-up of a large number of glioma patients. Our method comprises the following key steps. *First,* we construct both functional and structural brain networks. *Second,* we extract structural and functional connectomics features using diverse network metrics. *Third,* we propose a novel framework to effectively reduce the dimension of connectomics features by step-wisely selecting the most discriminative

features in a gradual, three-stage strategy. *Finally*, we use support vector machine (SVM) to predict the outcome.

2 Method

Figure 1 illustrates the proposed pipeline to automatically predict the survival time for HGG patients in three steps. In Sect. 2.1, we introduce the construction of brain networks based on the resting-state functional MRI (rs-fMRI) and diffusion tensor imaging (DTI). In Sect. 2.2, we describe how to calculate network properties based on graph theory using a binary graph and a weighted graph. As we add up clinical information, such as tumor location, size and histopathological types, we generate a long stacked feature vector. In Sect. 2.3, we propose a three-stage feature selection algorithm to remove redundant features. Finally, we apply an SVM on the selected features to predict the treatment outcome.

2.1 Brain Network Construction

Subjects—A total of 147 HGG patients were originally included in this study. We excluded patients lacking either rs-fMRI or DTI data. Patients with inadequate follow-up time, or died due to other reasons (e.g., road accident) were also excluded. Those with significant image artifacts and excessive head motion, as suggested by the following data processing, were also removed. All the images were checked by three experts to quantify the deformation of brain caused by tumor. Those with huge deformation, for which all three experts reached an agreement, were removed too. Finally, 34 patients who died within 650 days after surgery were labeled as "*bad*" outcome group, and the remaining 34 patients who survived more than 650 days after the surgery were classified into the "*good*" outcome group. The reason of using 650 days as a threshold is that the two-year survival rate for malignant glioma patients was reported to be 51.7 % [5]. We slightly adjusted the threshold to balance the sample size in the two groups.

Imaging—In addition to the conventional clinical imaging protocols, research-dedicated whole-brain rs-fMRI and DTI data were also collected preoperatively. The rs-fMRI has TR (repetition time) = 2 s, number of acquisitions = 240 (8 min), and a voxel size = $3.4 \times 3.4 \times 4$ mm^3. The DTI has 20 directions, voxel size = $2 \times 2 \times 2$ mm^3, and multiple acquisition = 2.

Clinical Treatment and Follow-up—All patients were treated according to clinical guideline for HGGs, including a total or sub-total resection of tumor entity during craniotomy and radio- and chemo-therapy after surgery. They were followed up in a scheduled time, e.g., 3, 6, 12, 24, 36, 48 months after discharging. Any vital event, such as death, was reported to us to let us calculate the overall survival time.

Image Processing—SPM8 and DPARSF [6] were used to preprocess rs-fMRI data and build functional brain networks. FSL and PANDA [7] were used to process the DTI data and build structural brain networks. Multimodal images were first co-registered within subject and then registered to the standard space. All these processes are following the commonly accepted pipeline and thus not detailed here.

Network Construction—For each subject, two types of brain networks were constructed (see descriptions below). For each network, we calculated graph theory-based properties from both binary and weighted graphs.

- **Structural Brain Network.** We parcellated each brain into 116 regions using Automated Anatomical Labeling (AAL) atlas, by warping the AAL template to each individual brain. The parcellated ROIs in each subject were used as graph nodes. The *weighted* network N_s^w can be constructed by calculating the structural connectivity strength $w_s(i,j) = \frac{2}{S_i + S_j} \sum_{i,j \varepsilon \mathrm{N}} l(f)$ for the edge connecting nodes i and j $(i,j \in N;\ i \quad j)$, where N is the set of all 116 nodes in the network, $l(f)$ represents the number of fibers linking each pair of the ROIs, and S_i denotes the cortical surface area of node i. The sum $S_i + S_j$ corrects the bias caused by different ROI sizes. The *binary* structural network N_s^b can be generated by setting the weight of the top 15 % edges to 1 after ranking the w_s descending, and the others to 0 [8].

- **Functional Brain Network.** Using the same parcellation, we extracted the mean BOLD time series $TS_i (i \in N)$ of each brain region. Then, we defined the functional connectivity strength $w_f(i,j)$ in the functional network by computing Pearson's correlations between two BOLD time series in each pair (i,j) of 116 brain regions: $w_f(i,j) = Corr(TS_i,\ TS_j)(i,j \in N;\ i \quad j)$, thus generating a weighted functional brain network N_f^w. The binary functional network N_f^b can be generated in the same way as described above.

2.2 Feature Extraction

Graph theory-based complex network analysis is used to independently extract multiple features from four networks $(N_s^w, N_s^b, N_f^w, N_f^b)$ for each subject. Since various graph metrics can reflect different organizational properties of the networks, we calculated four types of these metrics (i.e., degree, small-world properties, network efficiency properties, and nodal centrality) [9], which are detailed below.

- **Degree.** In each binary network, N_s^b and N_f^b, the node i's degree, k_i, counts the number of edges linked to it. In each of the weighted networks, N_s^w and N_f^w, the node degree is defined by $k_i = \sum_{j \in N, j \neq i} w_*(i,j)$, where * refers to s or f.

- **Small-world property.** This type of property is originally used to describe small-world, and can also be separately calculated for each node, including the clustering coefficient C_i (which measures local interconnectivity of the node i's neighbors) and the shortest path length L_i (which measures overall communication speed between node i and all other nodes). Specifically, in N_s^b and N_f^b, C_i is calculated through dividing the number of edges connecting i's neighbors by all possible edges linking i's neighbors (i.e., $k_i(k_i - 1)/2$). On the other hand, in N_s^w and N_f^w, C_i is calculated by a normalized sum of the mean

weight of two participating edges in all triangles with node i as a vertex. L_i is defined as the averaged minimum number of edges from node i to all other nodes in N_s^b or N_f^b, and the averaged minimum sum of weighted edges in N_s^w and N_f^w.

- **Network efficiency.** The efficiency property of a network measures how efficiently information is exchanged within a network, which gives a precise quantitative analysis of the networks' information flow. The global efficiency, $E_{global}(i)$, is defined as the sum of the inverse of the shortest path length between node i and all other nodes. The local efficiency, $E_{local}(i)$, represents the global efficiency of a subgraph, which consists of all node i's neighbors. The binary and weighted versions of shortest path length can result in binary and weighted efficiency metrics.

- **Nodal centrality.** Nodal centrality, B_i, quantifies how important of node i is in the network. A node with high B_i acts as a hub in the network. It is calculated as

$$B_i = \sum_{m \neq n \neq i \in N} \frac{L_{mn}(i)}{L_{mn}}$$, where L_{mn} is the total number of shortest paths from node m to node n, and $L_{mn}(i)$ is the number of these shortest paths passing through node i. Since $L_{mn}(i)$ and L_{mn} have both binary and weighted versions, B_i is calculated for each binary and weighted network.

These network metrics, which will be used as connectomics features, were computed as part of features using GRETNA [8]. We also add to them 13 clinical features (age, gender, tumor size, WHO grade, histopathological type, main location, epilepsy or not, specific location in all lobes, and hemisphere of tumor tissue). Therefore, a total of 2797 (6 metrics × 4 networks × 116 regions + 13 clinical features) features for each subject were used. The number of features is much greater than that of samples (68 subjects). This is quite troublesome for machine learning-based methods because of the overfitting problem and the interference from noise. Thus, we design a three-stage feature selection framework, as specified below, to select the most relevant features for our classification (i.e., prediction) problem.

2.3 Three-Stage Feature Selection

To identify a small number of features that are optimal for treatment outcome prediction, we propose a three-stage feature selection method to gradually select the most relevant features.

- **First stage.** We roughly select features that significantly distinguish the two outcome groups (i.e., "bad" and "good") using two sample t-tests with $p<0.05$.

- **Second stage.** RELIEFF [10] is used to rank the remaining features **X** and compute their weights. RELEFF is an algorithm, which estimates feature quality in classification. Many heuristic measures of feature quality usually suppose the independence of features, while actually they may be dependent. RELIEFF can correctly estimate the quality of each feature in classification problem with strong dependency assumption among features. The main idea of RELIEFF is to estimate how well each feature distinguishes itself from its neighbors that belong to other classes. Given a randomly selected feature R from the feature set **A**,

RELIEFF searches for its k-nearest neighbors first. Basically, it defines a cohort of neighbors as belonging to the same class of R (called nearest hit H), and also other neighbors as part of a different class (called nearest miss M). Then, it computes and updates the quality estimation W(**A**) for all features based on the distance from R to H and also distance from R to M. Therefore, the features can be descendingly ranked in **X** based on W(**A**).

- **Third stage.** A sequential backward selection [11] strategy was applied to carefully select a small group of significant features from **X**. Then, an inner SVM was wrapped into the feature selection framework to evaluate the predictive accuracy for candidate subset of features using a leave-one-out cross validation. The sequential backward selection is a feature selection strategy that sequentially removes one feature from back to front from **X**. The classification accuracy is recorded for the remaining subset of **X**. When no feature is left, the selection process stops and a subset of **X** with the highest classification accuracy is selected.

Next, the selected features are fed into an outer SVM with a leave-one-out cross validation to build the prediction model. To test which features are more useful for outcome prediction, we conducted five experiments, where different features were combined in different ways for classification (see Sect. 3).

3 Results

The outcome prediction accuracy of our proposed prediction framework is displayed in Table 1. Using only clinical features, the prediction accuracy only reaches 63.2 %. Notably, when using only the features from functional networks, the accuracy increases to 72 %. As we combine structural network features with functional ones, the classification rate reaches its apex (75 %, better than when only using clinical features). However, no improvement was noted when clinical features were further added, which means that the information contained in clinical features is somehow represented already in the brain functional and structural networks using graph theory. In order to test the results that we learned were random or not, we also did 30 times permutation test. The p-value of permutation test was 0.015 and the mean accuracy of 30 times permutation test was 49.1 %, which means that our results can reflect the intrinsic properties of the data to some degree. The most significant features shown in Table 2 (also drawn in Fig. 2) are those that were selected by our three-stage feature selection strategy more than 60 times out of 68 trials.

As reported in many previous studies, the most useful regions for HGG outcome prediction are highly correlated with movement, cognition, emotion, language and memory functions. The deteriorated structural and functional connections to these regions could influence the survival time. The most frequently selected ROIs from functional network are those in the cerebellum, which have dense functional connectivity to the neocortex and are closely associated with motor and cognitive functions. However, the most frequently selected ROIs from structural network are mostly located in the cortex and less overlapped with each other, which may indicate that the structural network is easily affected by brain tumors.

4 Conclusion and Future Works

In this paper, we have showed that complex brain network analysis, which is based on graph theory, is a powerful tool for treatment outcome prediction for high-glioma patients. Our findings highlighted the relevance of integrating functional and structural brain connectomics for HGG outcome prediction. Although the relationship between structural and functional brain networks is still poorly understood, our prediction framework remarkably benefitted from the use of brain connectomics for prognosis evaluation. In future works, we will incorporate the global graph properties (e.g., the averaged clustering coefficient, or network efficiency across all brain regions) as new features. In such case, individual heterogeneity of tumor characteristics can be better addressed. Also, more advanced graph metrics, e.g., assortativity, modularity, and rich-club value, can be taken into account for a more comprehensive network measurement. Moreover, intraoperatively derived features, e.g., extension of tumor resection, can also be integrated as important prognostic predictors.

Acknowledgments

References

1. Macyszyn L, Akbari H, et al. Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques. Neuro-oncology. 2015; 127

2. Jain R, Poisson LM, et al. Outcome prediction in patients with glioblastoma by using imaging, clinical, and genomic biomarkers: focus on the non-enhancing component of the tumor. Radiology. 2014; 272:484–493. [PubMed: 24646147]

3. Zacharaki EI, et al. Survival analysis of patients with high-grade gliomas based on data mining of imaging variables. AJNR. 2012; 33:1065–1071. [PubMed: 22322603]

4. Liu F, Guo W, et al. Multivariate classification of social anxiety disorder using whole brain functional connectivity. Brain Struct Funct. 2015; 220:101–115. [PubMed: 24072164]

5. Ostrom QT, et al. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2007–2011. Neuro-oncology. 2014; 16:vi1–vi63.

6. Yan C, Zang Y. DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. Front Syst Neurosci. 2010; 4:13. [PubMed: 20577591]

7. Cui Z, Zhong S, et al. PANDA: a pipeline toolbox for analyzing brain diffusion images. 2013

8. Wang J, Wang X, et al. GRETNA: a graph theoretical network analysis toolbox for imaging connectomics. Front Hum Neurosci. 2015; 9

9. Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. Neuroimage. 2010; 52:1059–1069. [PubMed: 19819337]

10. Robnik-Šikonja, M., Kononenko, I. An adaptation of Relief for attribute estimation in regression. ICML; 1997. p. 296-304.1997

11. Rückstieβ, T., Osendorfer, C., van der Smagt, P. Sequential feature selection for classification. In: Wang, D., Reynolds, M., editors. AI 2011 LNCS. Vol. 7106. Springer; Heidelberg: 2011. p. 132-141.
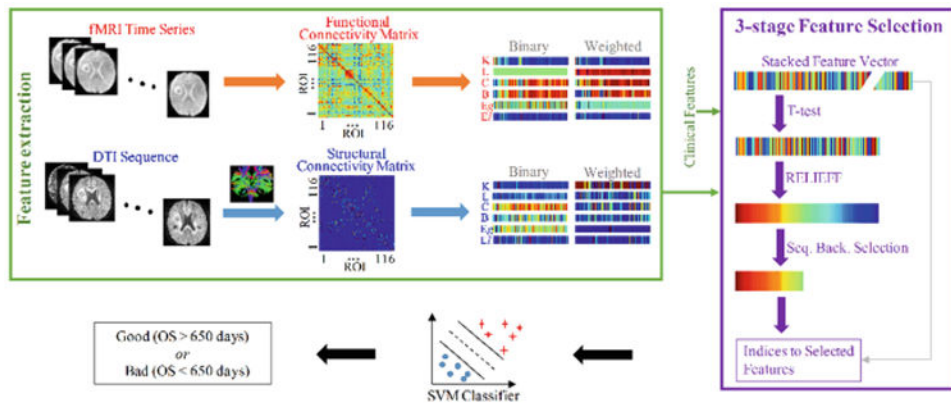
Fig. 1.
Proposed pipeline of treatment outcome prediction for high-grade glioma patients. (K: degree; L: shortest path length; C: clustering coefficient; B: betweenness centrality; Eg: global efficiency; E*l*: local efficiency; OS: overall survival. For details, please see Sect. 2.2).
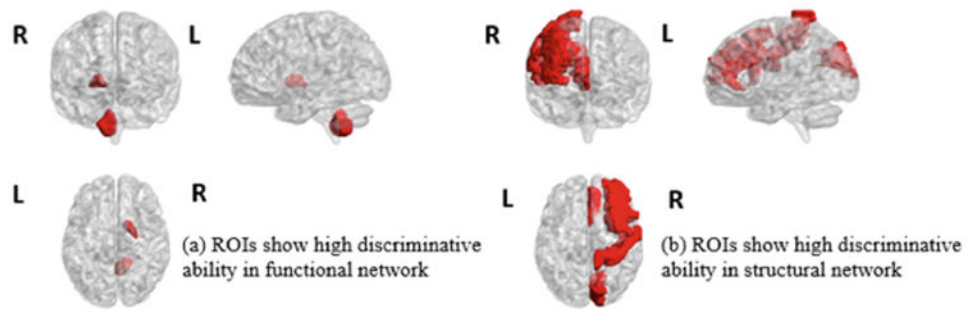
Fig. 2.
Discriminative ROIs with high predictive power in functional and structural brain networks, respectively.

Table 1

Prediction accuracy of using different sets of features.

Features	Accuracy (%)	Sensitivity (%)	Specificity (%)
Clinical infomation	63.2	61.8	64.7
Structural network	69.1	64.7	73.5
Functional network	72.1	70.6	73.5
Functional + Structural networks	**75**	**82.4**	67.6
Functional + Structural + Clinical	**75**	**82.4**	67.6

Table 2

The most useful features for outcome prediction.

Network Metrics	Predictive ROIs from fMRI	Predictive ROIs from DTI
Clustering coefficient	PAL R[a]	CUN R
Shortest path length	PAL R	IFGoper R
Global efficiency	CER9 R	MFG R, IFGoper R
Degree	CER9 R, PAL R	
Betweenness		ACG R, PoCG R

[a]For the full names of the brain regions, please see [4]. R: right side; L: left side.