Cerebral glioma grading using Bayesian Network with features extracted from multi-modality MRI

Jisu Hu #[1], Wenbo Wu #[2], Bin Zhu #[2], Huiting Wang[2], Renyuan Liu[2], Xin Zhang[2], Ming Li[2], Yongbo Yang[3], Jing Yan[4], Fengnan Niu[5], Chuanshuai Tian[2], Kun Wang[2], Haiping Yu[2], Weibo Chen[6], Suiren Wan*[1], Yu Sun*[1], and Bing Zhang*[2]

[1]*The Laboratory for Medical Electronics, School of Biological Sciences and Medical Engineering, Southeast University, Nanjing, China,* [2]*Department of Radiology, The Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, China,* [3]*Department of Neurosurgery, The Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, China,* [4]*Department of Oncology, The Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, China,* [5]*Department of Pathology, The Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, China,* [6]*Philips Healthcare, Shanghai, China*

*Corresponding authors, #these three authors contributed equally to this work.

TARGET AUDIENCE: This work will benefit neurologists and neuroradiologists who are in need of a tool helpful for glioma grading.

Purpose: Preoperative cerebral glioma grading is an important but difficult task in clinical practice. Although many state-of-the-art MR imaging techniques allow neuroradiologists to evaluate tumor grade in different aspects, there is still a need to combine findings from multiple modalities to make more reliable predictions of tumor grade. Bayesian Network [1](BN) is a powerful tool in artificial intelligence and has many applications in medical diagnosis. Hence, we aimed to develop a diagnosing tool for preoperative cerebral glioma grading based on BN that can more efficiently and reliably combine features extracted from multi-modality MR images, especially those from advanced MR imaging like perfusion weighted imaging and MR spectroscopic imaging (MRSI), and to evaluate the grading performance of the BN and individual features using real clinical data in a leave-one-out analysis.

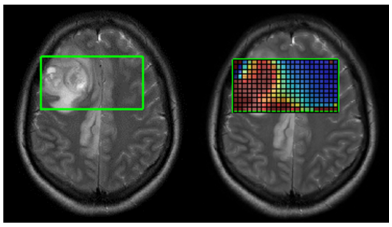

Fig. 1 The MRSI ROI projected onto the T2W image (left) and the pseudo-color map (right) of Cho/NAA where red regions indicate high Cho/NAA and blue ones indicate low Cho/NAA.

Methods: *Patients:* A total of 52 cerebral glioma patients, consisting of 30 high grade gliomas (WHO Ⅲ~Ⅳ) and 22 low grade (WHO Ⅰ~Ⅱ) ones, were performed MR scans preoperatively at a 3T MRI scanner (Achieva 3.0T TX, Philips Medical Systems, the Netherlands). Of the cases, all performed conventional MR imaging, 40 underwent PWI scans and 23 had MRSI scans. 30 cases were histopathologically graded and the rest were evaluated and assigned to the most probable grades by experienced neuroradiologists. *Feature extraction:* For each patient, the tumor volume was approximately calculated by the formula $V = \pi abc/6$, where *a* and *b* were the major and minor axes of the tumor in the axial T_2W image that exhibited the largest tumoral size and *c* the head-foot length in the coronal T_2W-FLAIR image that also presented the most of the tumor. Mass effect was measured by the maximum displacement of the mid-sagittal line from the T2W image with the heaviest mass effect. The T_1W contrast enhancement effect was categorized by negative, minor and severe contrast enhancement and was encoded as a discrete node in the BN. The ratios of rCBF and MTT values from tumoral regions referenced by normal regions respectively were selected for perfusion imaging features, and hence the perfusion feature node was 2-dimensional continuous. Metabolite concentrations detected from MRSI data were first quantitated using LCModel [2] and the maximal Cho/NAA across the ROI (Fig. 1) was selected as the spectroscopy feature for each case. *Statistical analysis:* The JMP software was used to perform Kruskal-Wallis test to see if there was statistical significance between high grade and low grade groups for each of the features extracted in the above methods. Additionally, Logistic regression analysis was performed on perfusion features. *BN construction and parameter learning:* The BN used in this work was built in cooperation of and confirmed by neuroradiologists. In Fig.2, rectangles denote discrete nodes and ellipses denote continuous nodes. Parameters of distributions for each node were learned after BN construction. Due to the fact that there were a number of patients who did not have PWI or MRSI scans, we used the EM (Expectation-Maximization) algorithm that can handle incomplete data to iteratively estimate the parameters. In each step of iteration, the algorithm contains two steps: the E-step computes a posterior distribution for each case using a chosen BN inference engine and the M-step gets the parameters that maximize the log-likelihood. Parameters are randomly initialized and the algorithm is running iteratively until either the parameters converge according to the criteria or the maximal number of iterations. In this study, we set maximum number of iteration to be 20 and the convergence threshold to be 0.001. *Probabilistic inference and grading:* For a new case of preoperative cerebral giloma, if we observe one or more of the five features, the posterior probability of tumor grade can be computed. The grade with the higher posterior probability was selected as the predictive grade of this case. Considering the comparatively small dataset, we adopted the leave-one-out analysis to evaluate the grading performance. In other words, we each time left one case out for testing and used the rest to learn the parameters.

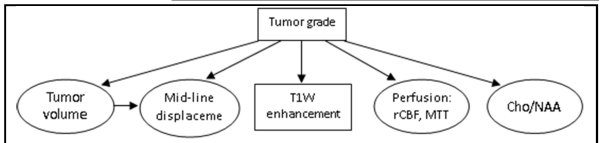

Fig.2 The network structure of the constructed BN

Results: Statistical analysis results showed significant differences ($p<0.05$) in all the features. Grading accuracies when different features observed are summarized in Table 1. It can be seen that with single observed feature, the feature of T_1W contrast enhancement came the first in the grading accuracy, followed by perfusion feature, Cho/NAA, mid-line displacement and tumor volume in that order. The overall accuracy with all the features observed is higher or at least equal to that with any single feature observed.

Table 1 Grading accuracies when different nodes observed.

Observed nodes	Accuracy (%)
Tumor volume	74.47
Mid-line displacement	76
T1W contrast enhancement	88.24
Perfusion feature: rCBV, MTT	87.5
Cho/NAA	87.5
All observed	88.24

Discussion: It can be seen that tumor volume and mid-line displacement are weaker in tumor grading and this is possibly due to that the data of high grade group overlap with those of low grade group. The much higher accuracies of the other three features indicate that they can possibly present more intrinsic characteristics of brain tumors. Particularly, the feature extraction methods for perfusion data and MRSI are different from many previous studies. In MRSI feature extraction, the metabolite concentrations are more reliable for that LCModel [2] treats a spectrum as a linear combination of all the basis metabolites detectable in a certain TE and estimates an optimal set of coefficients to fit the spectrum and minimize the residual signal. Moreover, the highest value of Cho/Cr generally appears in the region where cellular proliferation is most active while more neuron damaging often results in lower value of NAA/Cr. Hence, the maximal Cho/NAA across the ROI can represent the malignancy of a tumor. In perfusion feature selection, we chose rCBF and MTT instead of rCBV alone which was used in many previous studies and got comparable accuracy as Cho/NAA did. The grading system based on BN has several advantages. Firstly, each type of data can be described using different distributions and encoded as random variables in the network, which gives BN the ability of integrating different types of data. More importantly, the network structure represents the probabilistic casual relations and conditional dependencies in the nodes, which will make the results more interpretable. Secondly, BN is a tool for systematic inference under uncertainty and it can not only tell which grade the tumor is but also the confidence of that prediction as well. Thirdly, with EM algorithm, BN can deal with incomplete data, which is very important in clinical settings.

Conclusions: Results in this study suggest that BN is promising in combining features from multi-modal MRI for glioma grading. Considering the relative small dataset used here, the performance should be further validated with more data.

References: 1. Friedman, D.K.N., Probabilistic Graphical Models: Principles and Techniques. 2009. 2. Provencher, S.W., Estimation of Metabolite Concentrations from Localized In-vivo Proton NMR-Spectra. Magnetic Resonance in Medicine, 1993. 30(6): p. 672-679.