Research Article   Open Access

Quantitative Measurement of Oligodendrogliomas Histologica Features in Predicting the 1p/19q Co-Deletion Status

Yuchen Yang[1]*, Fuyong Xing[2] and Lin Yang[3]

[1]*Department of Statistics, College of Arts and Sciences, Lexington, KY, USA*
[2]*Department of Computer Science, College of Engineering, Lexington, KY, USA*
[3]*Division of Biomedical Informatics, Department of Biostatistics, College of Public Health, Department of Computer Science, College of Engineering, Lexington, KY, USA*

Abstract

Oligodendrogliomas are characterized by 1p/19q co-deletion, which generally correlates with subjective morphologic features like nuclei circularity ratio and texture features of the cancer nuclei. As cost control becomes more of an issue in medicine, upfront reflex molecular diagnostic testing for lesions like 1p/19q co-deletion may not be appropriate. This paper aims to develop a rigorous, unbiased digital imaging segmentation algorithm and statistical models that can accurately predict the likelihood of 1p/19q co-deletion based on morphology and texture, which would greatly improve cost-effectiveness in clinic trail. In this study, totally 28 gliomas of haematoxylin and eosin stained slides are comprised in this test cohort. Selected areas that had high tumour cell density were digitally analysed with a high-throughput image segmentation algorithm to automatically delineate the boundaries of the cell nuclei. Then we extracted the morphologic features and texture features based on the segmentation result, and applied them in to Lasso-logistic regression to build the correlation between these features with 1p/19q co-deletion status. As a comparison, we also used PAM (Prediction Analysis of Microarrays), RPA (Recursive Partitioning Analysis) to compare the predication performance. We find out that the circularity ratio of the cell, the variance of cell area for each patient, and parts of texture features effect the 1p/19q co-deletion status, and the false prediction rate of leave one out cross validation is at most around 10%. Moreover, we conduct survival analysis and find out two morphologic features and one texture features are significant influential to patients' survival time.

Introduction

Brain tumour is one of the most frequent cancers worldwide. An oligodendroglioma tumour is a slow growing brain tumour that is believed to originate from the oligodendrocytes of the brain or from a glial precursor cell. They occur primarily in adults (9.4% of all primary brain and central nervous system tumours) but are also found in children (4% of all primary brain tumours). The average age at diagnosis is 35 years [1]. These tumours are frequently located within the frontal, temporal or parietal lobes and cause seizures in a relatively high percentage of patients. Many oligodendrogliomas contain little specks of calcium (bone) and can easily. Based on both FISH and LOH findings [2], the present study suggests that the 1p/19q co-deletion in pure oligodendroglia tumours can be considered as a diagnostic, rather than a prognostic marker. 1p/19q genetic status has been examined in malignant gliomas since Cairncross et al. who described the clinical implications of 1p/19q co-deletion in patients with anaplastic [3-5]. Thus Oligodendrogliomas are characterized by 1p/19q co-deletion. Because oligodendrogliomas are less aggressive than their grade matched astrocytic counterparts, differentiating between an astrocytoma and an oligodendroglioma is a key component of surgical neuropathology. Unfortunately, such discrimination suffers from high inter-observer variability.

Furthermore, as cost control becomes more of an issue in medicine, upfront reflex molecular diagnostic testing for lesions like 1p/19q co-deletion may not be appropriate. An accurate, unbiased way to predict the likelihood of 1p/19q co-deletion would greatly improve cost-effectiveness, reserving this expensive test for cases in which co-deletion is reasonably possible. The recognition of molecular subsets among oligodendrogliomas has raised the question whether distinct mutations in associated genes may serve as prognostic markers (Figure 1). Recently, more and more research starts to focus on the

relationship between morphological features and oligodendrogliomas [6-9]. Partial oligodendroglial morphological features in GBMs were more frequently detected in tumours with 1p loss. In Ueki'study, which indicate that morphological features do necessary follow the genetic profile [10,11]. In that case; predict 1p/19q co-deletion status based on the morphologic features circularity ratio is desirable to achieve. From Scheie's study, there is a strong association between phenotype and genotype in oligodendroglial tumours. And even when all significant variables are accounted for, perfect prediction (100%) of 1p/19q status cannot be obtained [12]. So to find out a more accurate way to predict 1p/19q co-deletion status based on the morphologic features is a necessary. On the other hand, image texture features, which is a set of metrics calculated in image processing designed to quantify the perceived texture of an image, are still play an important role in image classification [13,14], as well as in prediction. However, to our knowledge, no study has documented a detailed method to mining the correlation between texture features and 1p/19q co-deletion status, to say nothing of, combine the morphologic features and texture features to predict 1p/19q co-deletion status.

In this paper, our first contribution is that we proposed a methodology that associated nuclei morphology which released from tissue slides images with the 1p/19q co-deletion status. Haematoxylin
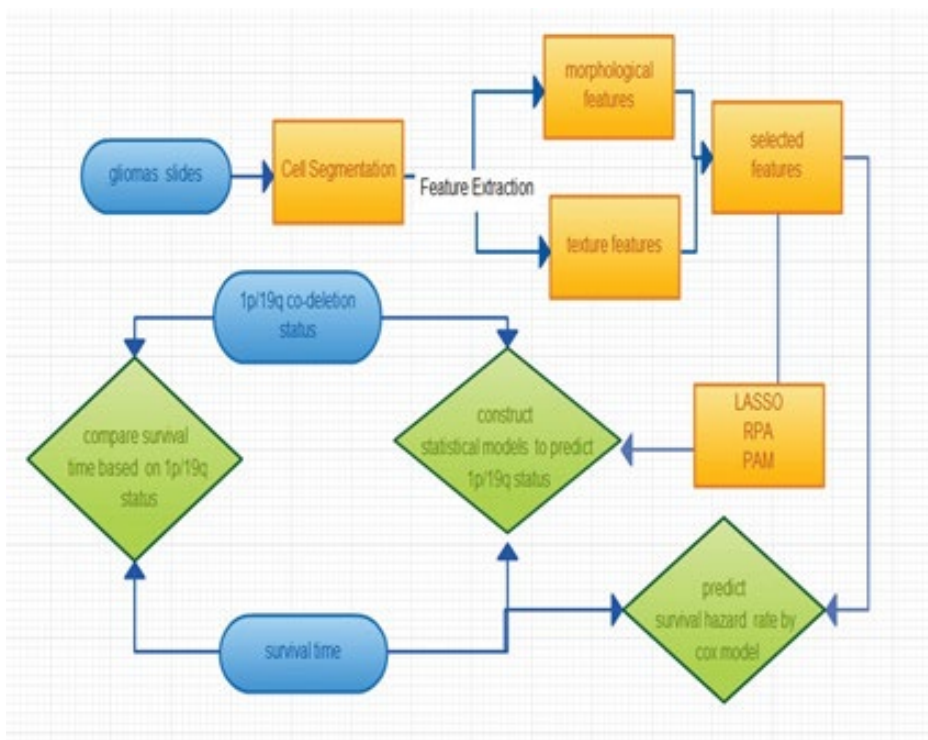
Figure 1: The flow chart of this study.

and eosin stained slides from 28 gliomas comprised the training cohort. Selected areas that had high tumour cell density (i.e., minimal non-neoplastic tissue contamination") were digitally analysed with a high-throughput image segmentation algorithm to automatically delineate the boundaries of the cell nuclei (typically hundreds in one image patch). Representative morphological features, such as the mean and variability of nuclear area, shape, circularity ratio, and perimeter and texture morphological features were collected based on segmentation result. Our second contribution is that we constructed three statistical models that can accurately predict the likelihood of 1p/19q co-deletion based on morphological features and texture features. Through the stepwise selection results for Lasso-logistic regression model significant image markers (variability of circularity ratio and two texture features) are selected. Based on these selected image markers, the Lasso model was constructed to predict the 1p/19q status for new cases. As comparison, PAM (Prediction Analysis of Microarrays) and RPA (Recursive Partitioning Analysis) were also used to select the significant image markers, trained and tested. At last, leave one out cross validation was used to compare the prediction performance (the false prediction rate) of Lasso with PAM and Recursive Partitioning Analysis. The accuracy of predicting 1p/19q status is significantly improved by our proposed models.

Furthermore, recent prospective randomized clinical trials have validated associations between combined 1p/19q co-deletion and prolonged overall survival of patients treated with radiation therapy with or without chemotherapy [3,4,15]. Just the same case like predicting 1p/19q co-deletion status, the work related to predict survival time and mining the hazard rate of patients and their morphologic features and texture features extracted from gliomas of haematoxylin and eosin stained slides are vacant in related research. In this paper, we also retrieve the previous study: analysis the difference of patients' survival time under different 1p/19q co-deletion status. Moreover, we built Cox proportional hazards model of patient and take extracted morphologic features and texture features as covariate. Two morphological features and one texture features are selected by forward stepwise selection. And we consider this as out third contribution.

Materials and Methods

Cell segmentation

A seed-controlled repulsivelevel setmethod [16,17] is applied to brain tumour cell segmentation in histopathology image.Since the number and the position of cells are not available, a prior, it is challenging to separate touching cells from each other. To this end, we employed a robust single-pass voting algorithm to accurately locate cell geometric centers, which are defined as seeds in this work. For each pixel (x,y) in image I(x,y), it defines a cone-shape voting area A with vertex at (x,y) and votes towards the direction of negative gradient based on the magnitude $\|\nabla I(x,y)\|$. To update the voting map V(x,y) with the same dimension as I(x,y), a Gaussian kernel $\kappa(u,v,\mu,\Sigma)$ is incorporated into the voting procedure [16]:

$$V(x,y) = V(x,y) + \sum_{(u,v)\in A} \|\nabla I(x,y)\| \kappa(u,v,\mu,\Sigma), \quad (1)$$

Where A is defined as a cone-shape region by the radial range (r_{min}, r_{max}) and the angular range

$$\Delta.\mu = \left(x + \frac{(r_{max}-r_{min})\cos\theta}{2}, y - \frac{(r_{max}-r_{min})\sin\theta}{2} \right)$$

denotes the mean of Gaussian kernel, and $\Sigma = \sigma^2 I_2$ represents the covariance matrix. The Gaussian kernel in (1) weights the magnitude for each pixel based on the distance between the pixel and the vertex. The closer it is to the cell center, the higher value it achieves. As a

sequence, the kernel encourages the voting toward the cell centre and the central pixels would finally obtain higher voting values compared with those near cell boundaries. Using a threshold to select these central pixels as seed candidates, mean shift [18] is applied to final seed detection by clustering the candidates. Since the candidates are always located in the cell central region, the final seeds would be correctly detected at the geometric centres.

With the detected seeds as warm initialization, a repulsive level set model [17] is employed to extract cell boundaries. Based on an interactive scheme, repulsive level set not only uses the competition of evolving contours to determine the membership of each pixel, but also applies the repulsion to prevent adjacent contours from overlapping. Let C_i (i=1,…,N) present the contours evolving toward the boundaries of N cells in image I, the level set energy function for cell segmentation with the interactive scheme can be expressed as follows:

$$E\left(C_1, C_2, \ldots, C_N\right) = \lambda_0 \sum_{i=1}^{N} \int_{in(C_i)} \left|I - c_i\right|^2 dxdy + \lambda_b \int_{\Omega_b} \left|I - c_b\right|^2 dxdy$$

$$+ \eta \sum_{i=1}^{N} \int_{0}^{1} g\left(\left|\nabla I\left(C_i\left(q\right)\right)\right|\right) \left|C_i'\left(q\right)\right| dq + \omega \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} A_i \cap A_j$$

(2)

Where A_i, i=1, N denotes the region closed by contour C_i and Ω_b represents the background. The in ( ) operator denotes the region inside cells. The c_i, c_b are the mean intensities of the cell region and background region, respectively. The λ_0, $\lambda 1$, η and ω are weights for the terms of cell region, background region, cell boundary, and repulsive scheme, respectively. Function g is chosen as a sigmoid function in the implementation with α representing the slope of the output curve and β representing the window size:

$$g\left(x\right) = \left(1 + e^{\left(\frac{x - \beta}{\alpha}\right)}\right)^{-1}, \qquad (3)$$

Cell segmentation is achieved by minimizing (2) using the level set framework. By introducing the Euler-Lagrange representation, equation (2) can be solved iteratively with the gradient descent method [17]. Due to the last term in (2), touching cells can be automatically and efficiently separated from each other (Figure 2).

Totally 28 gliomas of haematoxylin and eosin stained slides which cropped (select specific region of each pathology image to analysis) from the image provided by Norton Brain Tumour Centre, which are comprised in this test cohort. Each slide corresponds to a particular patient whose 1p/19q status is known. Selected areas that had high tumour cell density (i.e., minimal non-neoplastic tissue contamination) were digitally analysed with a high-throughput image segmentation algorithm to automatically delineate the boundaries of the cell nuclei (typically hundreds in one image patch).

Feature extraction

Based on the segmentation results, geometric features including nuclear area, perimeter, circularity index, and ratio between major and minor axes, are extracted for description. Due to the diversity among different cells within each image, the measurement of each feature would result in producing a unique distribution [19]. We apply the mean and the standard deviation to the geometrical features. Furthermore, one more robust texture feature, texton [20], is also calculated because of its strong discriminative power for classification [21]. In total, we extract 30 features consisting of 8 geometric features and one texton histogram containing 22 textons.
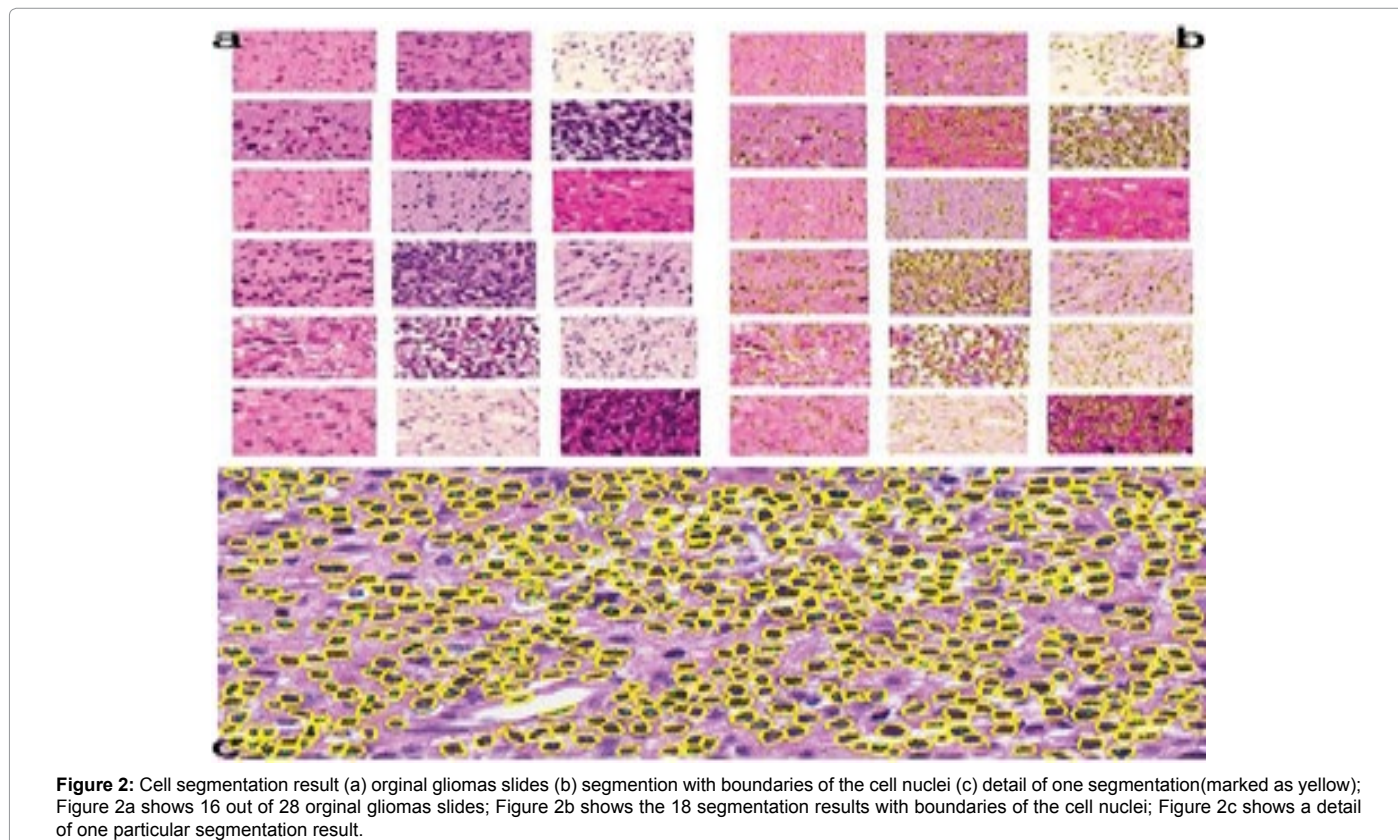

Figure 2: Cell segmentation result (a) orginal gliomas slides (b) segmention with boundaries of the cell nuclei (c) detail of one segmentation(marked as yellow); Figure 2a shows 16 out of 28 orginal gliomas slides; Figure 2b shows the 18 segmentation results with boundaries of the cell nuclei; Figure 2c shows a detail of one particular segmentation result.

Texton-based feature is one type of widely used texture features recently, and has achieved great successes in image segmentation [22], recognition [23], classification [24], etc. Textons are defined as the prototype filter response vectors, which are calculated by applying a filter bank to the images. In order to handle variation in cell size, intensity, and shape, the filter bank we used consists of 48 filters with 36 elongated filters at 6 orientations, 3 scales, and 2 phases, 8 center-surround difference of Gaussian filters, and 4 low-pass Gaussian filters. Therefore, each pixel is transformed into a 48-dimensional vector. Since the sliding windows are overlapping when we perform filtering, the responses will be overly redundant such that we can cluster them to form a compact representation. In the implementation, we use K-means for clustering (K is empirically set as 22). Each of the K-means centers (textons) encodes certain features so that similar pixels together with its neighbors should be mapped to the same class. The clustering center set actually forms a texton histogram which can be used as a texture feature for classification.

Statistical Models

Lasso-logistic regression

In traditional statistical analysis, we usually use ordinary least square (OLS) to obtain unbiased estimators, which is not satisfied due to the prediction accuracy and the difficulty in interpretation. To enhance the prediction accuracy, we sacrifice a little bias to reduce the variance by add penalty terms; to make it easy to interpret, we often would like to determine a smaller subset of among feature space that exhibits the strong effects. The lasso (shorted for least absolute shrinkage and select operator) which was originally proposed for linear regression models has become a popular model selection and shrinkage estimation method. It shrinks some coefficients and sets others to 0, and hence tries to retain the good features of both subset selection and ridge regression. In that case, lasso is a popular method for regression that uses an l_1 penalty to achieve a sparse solution [25,26]. The lasso estimate $(\hat{\alpha}, \hat{\beta})$ can be defined as follows:

$$\left(\hat{\alpha}, \hat{\beta}\right) = \arg\min\left\{\sum_{i=1}^{N}\left(y_i - \alpha - \sum_j \beta_j x_{ij}\right)^2\right\}$$

Subject to $\sum_j \left|\beta_j\right| \le t$ \hfill (4)

Where t is a tuning parameter, (y_i, x_{ij}) is the data set, i= 1,2,…,N and $\beta = (\beta_1, \beta_2, \dots \beta_p)^t$

In particular, when the response variable is binary, the linear logistic regression model is often used. Updating the y_i in equation (4) by $\log\left(\dfrac{\pi_i}{1-\pi_i}\right)$,

here $\pi_i = P\left(y_i = 1 \big| x_i, \beta\right) = \dfrac{1}{1+\exp\left(-\beta x_i\right)}$ \hfill (5)

In 2008, Friedman wrote a R package 'glmnet' to achieve a computational solution for lasso model, including lasso-logistic regression model (lasso two classification. The algorithms use cyclical coordinate descent, computed along a regularization path. The methods can handle large problems and can also deal efficiently with sparse features [27,28].

In our case, we want to predict 1p/19q co-deletion status more accurately by using a subset of representative features among all the features we extracted. Due to the sparse property of lasso, the lasso-

logistic regression model is a wise choice and can be applied to solve this problem. Using the lasso package: glmnet. We selected a 3-feature classifier among 30 features by stepwise selection: variability of circularity ratio, textons 9 and textons 12.

PAM (prediction analysis of microarrays)

PAM (Prediction Analysis of Microarrays) is a statistical technique for class prediction from gene expression data using nearest shrunken centroids. It is described in Tibshirani, Hastie, Narasimhan and Chu [26]. The method of nearest shrunken centroids identifies subsets of genes that best characterize each class. Shortly, we shrink the class centroids toward the overall centroids after standardizing by the within-class standard deviation for each gene. This standardization has the effect of giving higher weight to genes whose expression is stable within samples of the same class. Such standardization is inherent in other common statistical methods such as linear discriminant analysis.

Let $d_{ik} = \dfrac{\overline{x}_{ij} - \overline{x}_i}{m_k\left(s_i + s_0\right)}$ \hfill (6)

Where x_{ij} be the expression for genes i=1,2,…p, samples j=1,2,…n, K is the number of class, and $m_k = \sqrt{1/\left(n_k + \dfrac{1}{n}\right)}$ makes the $m_k\, s_j$ equal to the estimated standard error of the numerator in d_{ik}. In the denominator, the value s_0 is a positive constant (with the same value for all genes), included to guard against the possibility of large d_{ik} values arising by chance from genes with low expression levels. We set s_0 equal to the median value of the s_i over the set of genes. Then each d_{ik} is shank toward zero, giving d_{ik}' and yielding shrunken centroids or prototypes d_{ik}, where

$$d_{ik}' = \text{sign}\left(d_{ik}\right)\left(\left|d_{ik}\right| - \Delta\right)_+ \hspace{1cm} (7)$$

+ means positive part ($t_+ = t$ if t > 0 and zero otherwise). This method usually produces more reliable estimates of the true means.

The technique is general and can be used in many other classification problems. It can also be applied to image classification. PAM Software for the R package [29] has been available for some time now.

Using PAM, we selected a 7-feature among 30 features: they are mean of circularity ratio, variability of circularity ratio, variability of nuclear area, textons 9, textons 12, textons 6 and textons 3.

RPA (recursive partitioning analysis)

Recursive partitioning methods have become popular and widely used for non-parametric regression and classification in many fields. Especially random forests, which can deal with large numbers of predictor variables even in the presence of complex interactions, have been applied successfully in genetics, clinical medicine and bioinformatics within the past few years.

Classification and regression trees are a simple nonparametric regression approach. Their main characteristic is that the feature space, the space spanned by all predictor variables, is recursively partitioned into a set of rectangular areas, as illustrated below. The partition is created such that observations with similar response values are grouped. After the partition is completed, a constant value of the response variable is predicted within each area [30].

There is a R package "rpart" available to achieve this method for classification [31]. As a comparison of lasso model, we use PAM and RPA. We selected a 6-feature among 30 features: they are mean of

circularity ratio, variability of circularity ratio, textons 3, textons 6, textons 9 and textons 12.

Cox proportional hazards model

Proportional hazards models are a class of survival models, which states that covariates are multiplicatively related to the hazard. Sir David Cox observed that if the proportional hazards assumption holds (or, is assumed to hold) then it is possible to estimate the effect parameter without any consideration of the hazard function.

Result

Based on the representative morphological features we computed from segmentation result and texture feature cropped from original gliomas of haematoxylin and eosin stained slides, and applied them into our three statistical classification models, we get the following feature selection result.

As expected, only parts of features out of 30 total features indeed are highly significantly correlated with 1p/19q co-deletion (Table 1).

From Table 1, we can find variability of circularity ratio; textons 9 and textons 12 are significant feature in all the three model, which may play an most important role in predicting 1p/19q co-deletion.

Variability of circularity ratio, textons 3 and textons 6 occur two times in the three statistical models, which may contribute secondarily to the performance of 1p/19q co-deletion prediction.

Figure 3a shows the stepwise features selection result for lasso-logistic model. Clearly, variability of circularity ratio, textons 9 and textons 12 are chosen under that method. Figure 3b shows the model deviance by choosing different value of penalty parameters. The left vertical curve is the penalty parameter chosen by minimizing the model deviance, while the right vertical curve is the penalty parameter chosen by minimizing the 1 standard error of deviance. In practical, we prefer the 1 standard error criteria since it reflect more variability [27] (Figure 3).

Figure 4a shows features selection result for PAM and Figure 4b shows the performance of each selected feature in predicting 1p/19q co-deletion status (Figure 4).

Leave one out cross validation

Cross validation, sometimes called rotation estimation, is a technique for assessing how the results of a statistical analysis will generalize to an independent data set. It is mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice. One round of cross validation involves partitioning a sample of data into complementary subsets, performing the analysis on one subset (called the training set), and

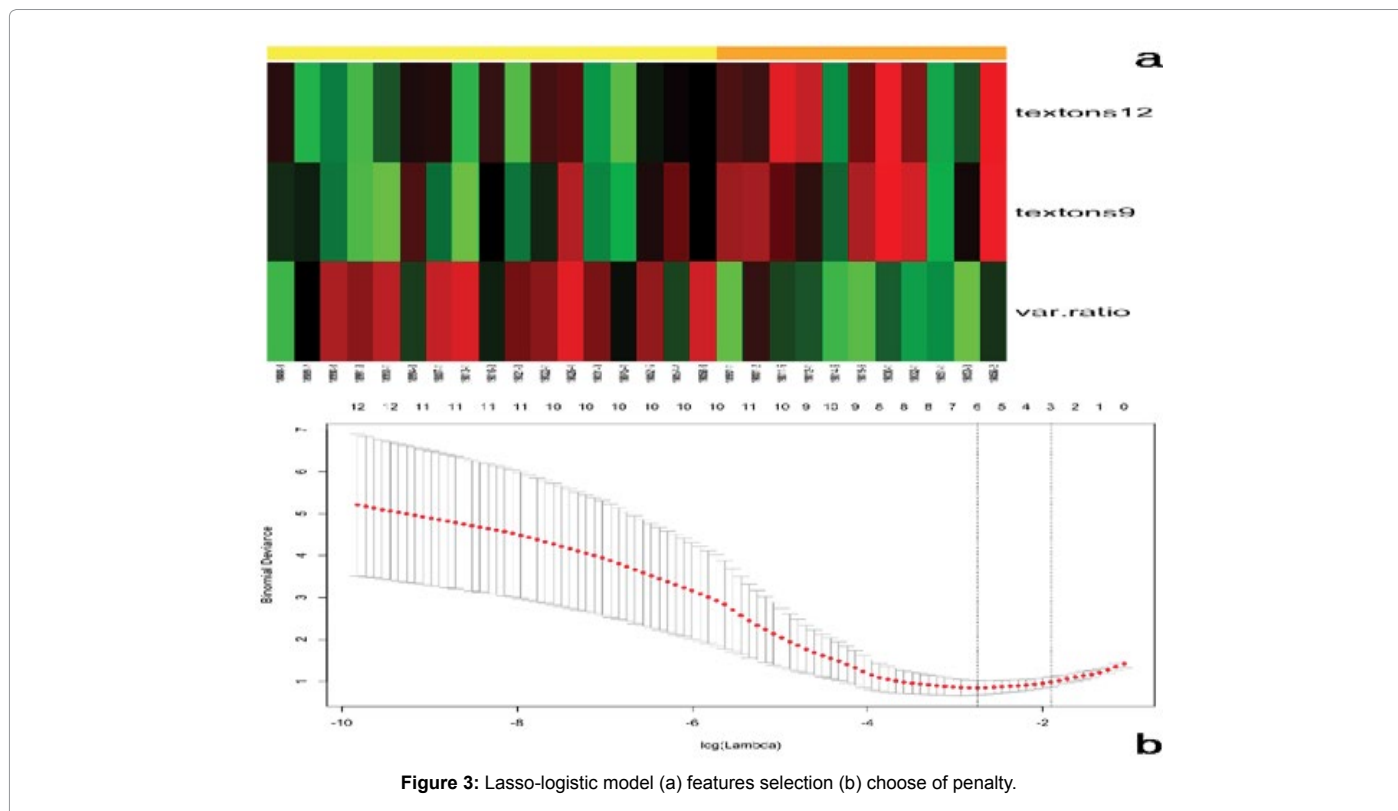| Method	M.CR	V.CR	V.A	Textons3	Textons6	Texstons9	Textons12
Lasso		Y				Y	Y
PAM	Y	Y	Y	Y	Y	Y	Y
RPA	Y	Y		Y	Y	Y	Y
M.CR: Mean of Circularity Ratio; V.CR: Variability of Circularity Ratio; V.A: Variability of Nuclear Area.							

Table 1: Feature selection result.

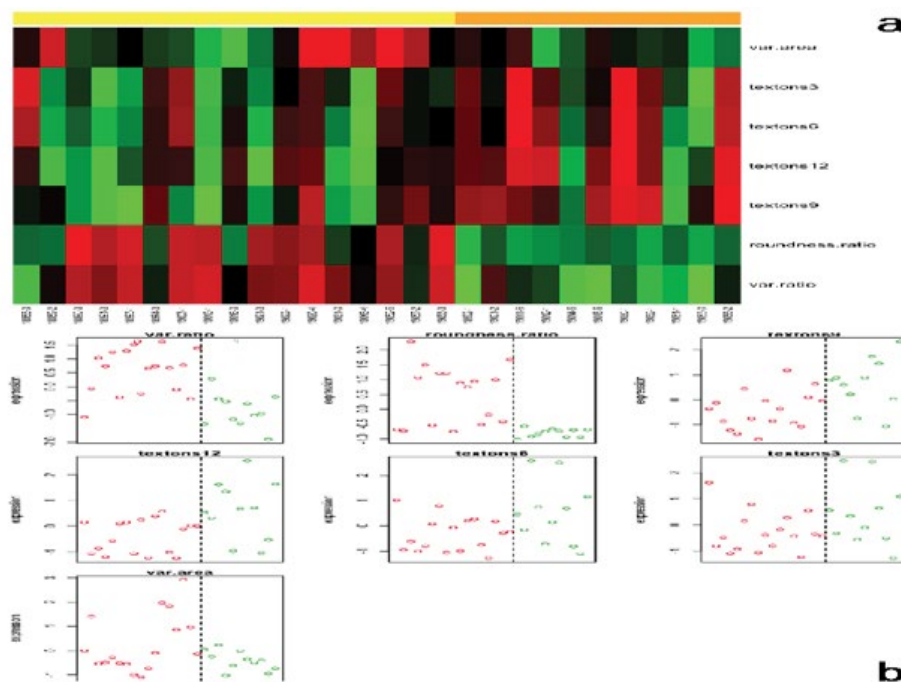**Figure 3:** Lasso-logistic model (a) features selection (b) choose of penalty.

Figure 4: PAM (a) features selection (b) performance of each selected feature.

validating the analysis on the other subset (called the validation set or testing set). To reduce variability, multiple rounds of cross-validation are performed using different partitions, and the validation results are averaged over the rounds [32].

Leave-one-out cross-validation (LOOCV) is widely used when the data size we have is limited. As the name suggests, it involves using a single observation from the original sample as the validation data, and the remaining observations as the training data. This is repeated such that each observation in the sample is used once as the validation data. Leave-one-out cross-validation is computationally expensive because it requires many repetitions of training.

As we only collect the information from 28 patients, LOOCV was used to compare the prediction performance of Lasso, PAM and RPA.

We find out that the false prediction rates are 7.14%, 17.85% and 10.71%, respectively, for Lasso, PAM and RPA.

Compared to PAM and RPA, Lasso-logistic regression model achieved least false prediction rate (7.14%), while PAM which include the most features have a worst prediction performance. On the other hand, RPA model have the best prediction performance result when the patient has 1p/19q co-deletion (false prediction rate is 0). To sum up, we declare that variability of nuclear area is not as much as important to other 6 features in Table 1; mean of circularity ratio, variability of circularity ratio, textons 3, textons 6 are more important features when the patient has 1p loss then not.

Survival analysis

To further study, we conduct survival analysis of patients' survival time to 30 morphologic. In the group with 19p loss, a significantly better (p = 0.022) survival period was found for the patients (median was three times longer than those with 1p loss). In that case, we'd like

to believe that 1p/19q co-deletion status is a critical factor to influence the survival time for brain tumour patients. Since we have already built the relationship between morphologic features and texture features and 1p/19q co-deletion status, to construct cox proportional hazards model of patient and take extracted morphologic features and texture features as covariate is appropriate. Three significant features (mean of circularity ratio, variability of circularity ratio and textons 3) are selected by forward stepwise selection method (Figure 5).

From that we can see mean of circularity ratio (p = 0.0048), variability of circularity ratio (0.0049) and textons 3 (p = 0.0207) have the parameter estimate 0.0968, 0.0121 and 0.024 respectively. It means the survival hazard rate will 0.0968, 0.0121 and 0.024 times higher than the baseline hazard rate we pre-choose for any unit increase in mean of circularity ratio, variability of circularity ratio and textons 3.

Discussion and Conclusion

The major aim of the study is to investigate the correlation between nuclei morphology, nuclei texture features and 1p/19q co-deletion status based on our proposed segmentation methodology. Based on a cohort of 28 patients, we have achieved at most 93% prediction accuracy through three statistical model, which confirms the long-standing observation that nuclear morphology (variability of circularity ratio) and parts texture features (textons 9 and 12) are highly correlated with 1p/19q co-deletion status. The Lasso-logistic regression model produces the best prediction results using these discovered image markers. Furthermore, survival analyses are conducted based on patients' survival time to those image markers and 1p/19q co-deletion status respectively. We also detect the significant difference of survival time with or without 1p loss, and cox model is used to select those image markers regarding to survival time. Mean of circularity ratio, variability of circularity ratio and textons 3 are selected, which means they are influential to survival time. However, our study has two
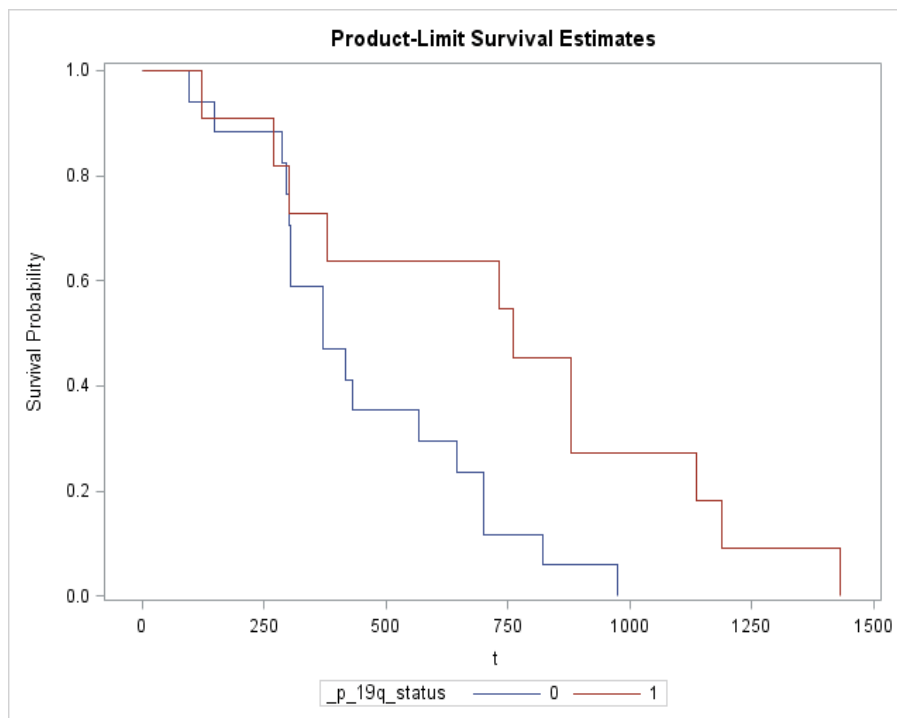
Figure 5: Cumulative survival rates for patients with different 1p/19q co-deletion status.

limitations. Firstly, the morphologic features we chose is subjective, and there might be possible we have ignored some other influential morphologic features. Secondly, we cropped/selected specific region of each pathology image to analysis instead of the whole image region. In that case, some critical latent nuclei might be ignored. In the future work, the additional morphologic features are currently being explored to further enhance prediction accuracy. Once a segmentation algorithm is sufficiently optimized, a validation cohort from Norton Brain Tumour Centre will be analysed. Additionally, instead of calculating from select regions, we also plan to apply the automatic image analysis on the whole slide scanned pathology images.

Acknowledgement

References

1. Huang CI, Chiou WH, Ho DM (1987) Oligodendroglioma occurring after radiation therapy for pituitary adenoma. J Neurol Neurosurg Psychiatry 50: 1619-1624.

2. Gadji M, Fortin D, Tsanaclis AM, Drouin R (2009) Is the 1p/19q deletion a diagnostic marker of oligodendrogliomas?. Cancer Genet Cytogenet 194: 12-22.

3. Cairncross G, Berkey B, Shaw E, Jenkins R, Scheithauer B, et al. (2006) Phase III trial of chemotherapy plus radiotherapy compared with radiotherapy alone for pure and mixed anaplastic oligodendroglioma: Intergroup Radiation Therapy Oncology Group Trial 9402. Journal of clinical oncology 24: 2707-2714.

4. Cairncross JG, Ueki K, Zlatescu MC, Lisle DK, Finkelstein DM, et al. (1998) Specific genetic predictors of chemotherapeutic response and survival in patients with anaplastic oligodendrogliomas. Journal of the National Cancer Institute 90: 1473-1479.

5. Cairncross G, Jenkins R (2008) Gliomas with 1p/19q codeletion: a.k.a. oligodendroglioma. Cancer J 14: 352-357.

6. Bigner SH, Rasheed BK, Wiltshire R, McLendon RE (1999) Morphologic and molecular genetic aspects of oligodendroglial neoplasms. Neuro Oncol 1: 52-60.

7. Buccoliero AM, Castiglione F, Degl'innocenti DR, Moncini D, Paglierani M, et al. (2012) Glioneuronal tumor with neuropil-like islands: clinical, morphologic, immunohistochemical, and molecular features of three pediatric cases. Pediatric and developmental pathology: the official journal of the Society for Pediatric Pathology and the Paediatric Pathology Society 15: 352-360.

8. Ronin VS (1990) [Morphologic characteristics of tumor cells in the cystic fluid of oligodendroglioblastoma]. Lab Delo : 79-80.

9. Gupta M, Djalilvand A, Brat DJ (2005) Clarifying the diffuse gliomas: an update on the morphologic features and markers that discriminate oligodendroglioma from astrocytoma. Am J Clin Pathol 124: 755-768.

10. Ueki K, Nishikawa R, Nakazato Y, Hirose T, Hirato J, et al. (2002) Correlation of histology and molecular genetic analysis of 1p, 19q, 10q, TP53, EGFR, CDK4, and CDKN2A in 91 astrocytic and oligodendroglial tumors. Clinical cancer research: an official journal of the American Association for Cancer Research 8: 196-201.

11. Mizoguchi M, Yoshimoto K, Ma X, Guan Y, Hata N, et al. (2012) Molecular characteristics of glioblastoma with 1p/19q co-deletion. Brain Tumor Pathol 29: 148-153.

12. Scheie D, Cvancarova M, Mørk S, Skullerud K, Andresen PA, et al. (2008) Can morphology predict 1p/19q loss in oligodendroglial tumours? Histopathology 53: 578-587.

13. Robert M, Haralick KS, Itshak D (1973) Textural Features For Image Classification. Transactions On Systems SMC-3: 610-621.

14. Herlidou-Même S, Constans JM, Carsin B, Olivie D, Eliat PA, et al. (2003) MRI texture analysis on texture test objects, normal brain and intracranial tumors. Magn Reson Imaging 21: 989-993.

15. Lindegaard KF, Mørk SJ, Eide GE, Halvorsen TB, Hatlevoll R, et al. (1987) Statistical analysis of clinicopathological features, radiotherapy, and survival in 170 cases of oligodendroglioma. J Neurosurg 67: 224-230.

16. Qi X, Xing F, Foran DJ, Lin Yang (2012) Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set. IEEE Transactions on Biomedical Engineering (TBME) 59: 754-765.

17. Yan P, Zhou X, Shah M, Wong ST (2008) Automatic segmentation of high-throughput RNAi fluorescent cellular images. IEEE Trans Inf Technol Biomed 12: 109-117.

18. Comniciu D, Meer P (2002) Mean shift: a robsut approach toward feature space analysis. IEEE transactions on pattern analysis and machine intelligence (PAMI) 24: 603-619.

19. Gao L, Zhou H, Thrall MJ, Li F, Yang Y, (2011) Label-free highresolution imaging of prostate glands and cavernous nerves using coherent anti-Stokes Raman scattering microscopy. Biomedical Optics Express 2: 915-926.

20. Leung T, Malik J (2001) Representing and recognizing the visual appearance of materials using three-dimensional textons. International Journal of Computer Visision (IJCV) 43: 29-44.

21. Qi X, Kim H, Xing F, Parashar M, Foran DJ, et al. (2012) The analysis of image feature robustness using cometcloud. J Pathol Inform 3: 33.

22. Malik J, Belongie S, Shi J, Leung T (1998) Textons, contours, and regions: cue integration in image segmentation. IEEE International Conference on Computer Vision (ICCV) 1999: 918-925.

23. Cula OG, Dana KJ (2004) 3D texture recognition using bidirectional feature histograms. International Journal of Computer Visision (IJCV) 59: 33-60.

24. Varma M, Zisserman A (2009) A statistical approach to material classification using image path exemplars. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 31: 2032-2047.

25. Tibshirani R (1996) Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B (Methodological) 58: 267-288.

26. Tibshirani R, Hastie T, Narasimhan B, Chu G (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci USA 99: 6567-6572.

27. Meier L, Van De GS, Bühlmann P (2008) The group lasso for logistic regression. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 70: 53-71.

28. Friedman J, Hastie T, Tibshirani R (2010) Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw 33: 1-22.

29. Ihaka R, Gentleman RR (1996) A language for data analysis and graphics. Journal of Computational and Graphical Statistics 5: 299-314.

30. Gaspar LE, Scott C, Murray K, Curran W (2000) Validation of the RTOG recursive partitioning analysis (RPA) classification for brain metastases. International Journal of Radiation Oncology, Biology and Physics 47: 1001-1006.

31. Therneau TM, Atkinson B, Ripley B (2010) rpart: Recursive Partitioning. R package version 3.1-46. Ported to R by Brian Ripley http://CRAN R-project org/package=rpart 3:2011.

32. Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. IJCAI'95 Proceedings of the 14th international joint conference on artificial intelligence 2: 1137-1143.