Age groups related glioblastoma study based on radiomics approach

Zeju Li\textsuperscript{a}, Yuanyuan Wang\textsuperscript{b,}\textsuperscript{a}, Jinhua Yu\textsuperscript{a,}\textsuperscript{b}, Yi Guo\textsuperscript{a,}\textsuperscript{b} and Qi Zhang\textsuperscript{c}

\textsuperscript{a}Department of Electronic Engineering, Fudan University, Shanghai, China; \textsuperscript{b}Key laboratory of Medical Imaging Computing and Computer Assisted Intervention of Shanghai, Shanghai, China; \textsuperscript{c}School of Communication and Information Engineering, Institute of Biomedical Engineering, Shanghai University, Shanghai, China

\section*{ABSTRACT}
Glioblastoma is the most aggressive malignant brain tumor with poor prognosis. Radiomics is a newly emerging and promising technique to reveal the complex relationships between high-throughput medical image features and deep information of disease including pathology, biomarkers and genomics. An approach was developed to investigate the internal relationship between magnetic resonance imaging (MRI) features and the age-related origins of glioblastomas based on a quantitative radiomics method. A fully automatic image segmentation method was applied to segment the tumor regions from three dimensional MRI images. 555 features were then extracted from the image data. By analyzing large numbers of quantitative image features, some predictive and prognostic information could be obtained by the radiomics approach. 96 patients diagnosed with glioblastoma pathologically have been divided into two age groups (<45 and $\geq$45 years old). As expected, there are 101 features showing the consistency with the age groups (T test, $p<.05$), and unsupervised clustering results of those features also show coherence with the age difference (T test, $p=.006$). In conclusion, glioblastoma in different age groups present different radiomics-feature patterns with statistical significance, which indicates that glioblastoma in different age groups should have different pathologic, protein, or genic origins.

\section*{KEYWORDS}
Word; radiomics; glioblastoma; age; prognosis; heat map; magnetic resonance images

\section*{1. Introduction}
Glioma is the most common malignant brain tumor, categorized into astrocytoma, oligodendroglioma, oligoastrocytoma and glioblastoma histologically, with grade I to IV according to tumor malignancy, according to WHO criteria \cite{1,2}. Glioblastoma is the most lethal subtype with dismal prognosis. Patients only have an average survival time of 12–15 months even under aggressive treatment \cite{3}. Nowadays, prediction of patient prognosis depends on not only pathological diagnosis, but also clinical parameters and genomic heterogeneity which served as crucial biomarkers to customized effective treatment paradigm. Prognosis estimation of high-grade gliomas is always considered to be a very difficult task because of rapid pre- and post-operative disease evolution \cite{2}. Effective prediction of survival factor is very beneficial to the treatment of glioma patients \cite{4}.

It has been noticed that age is an important marker to predict patient outcome with glioblastoma. It was found to be significantly associated with different types of glioma in a large cohort survey by Yang et al. \cite{5}. Li et al. revealed that "age $\leq$50 years old" was an independent favorable factor for glioblastoma by using cox proportional hazards modeling \cite{6}. In high-grade (grade III and grade IV) gliomas, age was also confirmed to be a prognostic factor by Gregorio et al. \cite{7}. However, how age influences patients' outcome is still under investigation. Nowadays, several papers based on large sample size found out the fact that glioma patients in different age groups may have distinct genetic background which provided explanation to their different tumor growth patterns and clinical outcomes \cite{8}. However, still very few studies have been focused on the relationship between age and MRI features in glioblastoma patients. The correlation between tumor location and age were demonstrated by Y. Wang et al. \cite{9}. An age related MRI study based on edema was carried out by C. Seidel et al. \cite{10}, while characteristics other than edema were not considered in that article.

In order to extract prognostic information from medical images which is not visible to human eyes, a
radiomics method has been adopted. Radiomics method is referred as a series of approaches including the extraction of plenty of relatively high throughput information from medical images and subsequent correlation analysis [11]. As an emerging field, radiomics has been successfully utilized in the analysis of genetic and clinical information (such as age and gender) in a non-invasive way for many cancer types, for example, lung cancer, head-and-neck cancer et al. [12–14].

In the following section, patients MRI images were analyzed and the relationship between MRI images and patients age was explored using radiomics method. Features extracted from MRI images showed statistical distinction between different age groups. The results in this paper demonstrated that glioblastoma patients of different age showed different characteristics in MRI images and indicated that different glioblastoma patients may be regarded as different disease subtype.

2. Material and method

In this paper, a fast and fully automatic segmentation method developed by our group has been used to get the tumor regions in MRI images. The method consisted of three steps. Firstly, a three dimensional (3D) bounding box has been used to find the volume of interest (VOI). Secondly, a reflecting symmetry detection method has been used to verify the validity of the VOI. Finally, with the accurate initial VOI, the automatic 3D grow-cut method has been utilized to provide accurate segmentation of gliomas.

After the tumor image segmentation, a radiomics approach has been developed. 555 features have been extracted from 96 glioblastoma patients T1 Contrast MRI images including 21 intensity features, 15 shape features, 39 texture features and 480 wavelet features. By following a previous study by Lacroix et al. [15], 96 patients have been divided into two groups according to the threshold of 45 years old. 101 most remarkable features consistent with the age group through the students’ tests (T test) have been selected, of which p values are less than .05. 101 distinct radiomics features have been taken to draw a heat map and we compared the age groups with the cluster results. Eventually, significant distinguish between different age groups by the distinction of radiomics features has been found.

2.1 Patients

All patients enrolled in our study were newly diagnosed and untreated, primary glioblastomas who underwent surgical resection in Shanghai Huashan Hospital during September 2013 to July 2015. We retrieved the MRI images from Huashan Glioma Image Database on the basis of two criteria: 1) clear and recognizable T1 Contrast MRI images, 2) absence of severe edema. The baseline data of all 96 patients are presented in Table 1, separated into two groups according to the age division. This experiment was approved by Huashan Hospital Ethics Committee and informed consent was obtained from every patient.

| Table 1. Patients characteristics. |
| Characteristic | n   | Percentage |
|----------------|-----|------------|
| Gender         |     |            |
| Male           | 59  | 64.5%      |
| Female         | 37  | 38.5%      |
| Age group (range from 27 to 78 years old) | | |
| <45 years old  | 29  | 30.2%      |
| ≥45 years old  | 67  | 69.8%      |

2.2 Tumor segmentation

In this study, in order to gain the objective and precise segmentation of the tumor regions in MRI images, a fully automatic method for segmenting gliomas in 3D MRI images has been realized. The 2D bounding box method has been extended to 3D to get the rough VOI. And so as to make up the disadvantage of the bounding box which cannot detect tumor regions across the mid-sagittal plane, a reflecting symmetry method has been proposed to get the correct VOI. At last, the grow-cut method has been extended to an automatic 3D grow-cut method to achieve a precise boundary of the tumor regions.

The segmentation method has been proved to be not only more convenient and feasible than the manual method but also faster than the 2D homogeneous segmentation method both in practice and theoretical analysis. What is more important, it succeeds in providing objective and accurate results for all 96 gliomas data. All segmentation results were proved by two experienced neuroradiologists. Taking the manual segmentation results as the ground truth, the Dice similarity coefficient [16] of the segmentation results were around 0.80 for each case.

2.3 Tumor segmentation

According to radiomics, 555 high throughput image features have been extracted from the region of segmented gliomas in all glioblastoma T1 Contrast MRI images [12,17].

The lists of total 555 features are concisely summed up in Table 2. The customized three dimensional features include intensity features, shape features, texture features.
features and wavelet decomposition features. All features have been extracted from MRI images in a raw data form. A brief description of these features is described here.

1) Intensity features: intensity features reflect histogram distribution of images, the histograms of images have been divided into 100 levels, and the gray-level statistics characteristics of original MRI images and those histograms such as kurtosis, skewness and variance have been acquired.

2) Shape features: shape features describe the morphological structure of the tumor regions, reflecting the size, smoothness, saturation and other characteristics of morphology. Two dimensional features such like Fourier-descriptors are extended to three dimensions by applying area weighted method of each piece.

| Feature name                  | n     |
|-------------------------------|-------|
| Intensity                     | 21    |
| T-energy                      |       |
| T-kurtosis                    |       |
| T-mean absolute deviation     |       |
| T-media                       |       |
| T-range                       |       |
| T-skewness                    |       |
| H-uniformity                  |       |
| Gauss-fitting                 |       |
| Gauss-fitting                 |       |
| H-variance                    |       |
| H-kurtosis                    |       |
| Shape                         | 15    |
| Compactness                   |       |
| Max-length                    |       |
| Sphericity                    |       |
| Surface to volume ratio       |       |
| Region to bounding-box ratio  |       |
| Min minor-length              |       |
| Orientation                   |       |
| Fourier-descriptors           |       |
| Texture                       |       |
| GLCMa                         | 8     |
| Energy                        |       |
| Correlation                   |       |
| Variance                      |       |
| Entropy                       |       |
| GLRMLMa                       | 13    |
| Short run emphasis            |       |
| Gray-level nonuniformity      |       |
| Run percentage                |       |
| High gray-level run emphasis  |       |
| Short run high gray-level emphasis |     |
| Long run high gray-level emphasis |   |
| Run-length variance           |       |
| GLSZMc                        | 13    |
| Small zone emphasis           |       |
| Gray-level nonuniformity      |       |
| Zone percentage               |       |
| High gray-level zone emphasis |       |
| Small zone high gray-level emphasis |     |
| Large zone high gray-level emphasis |   |
| Zone-size variance            |       |
| NGTDMd                        | 5     |
| Coarseness                    |       |
| Busyness                      |       |
| Strength                      |       |
| Wavelet decomposition         | 480   |
| LLL decomposition             |       |
| LHL decomposition             |       |
| LHH decomposition             |       |
| HHH decomposition             |       |
| Total of customized features  | 555   |

---

*a GLCM: Gray-level co-occurrence matrix.
*b GLRLM: Gray-level run-length matrix.
*c GLSZM: Gray-level size zone matrix.
*d NGTDM: Neighborhood gray-tone difference matrix.
3) Texture features: texture features contain the homogeneity level of images, four kinds of effective and valid high-order texture features have been selected for the analysis. In order to get more reliable results about MRI image texture, voxels within the tumor regions with intensities outside the range $\mu \pm 3\sigma$ were excluded ($\mu$ stands for the mean value and $\sigma$ stands for the standard deviation), as suggested by Collewet et al. [18].

a) GLCM: GLCM contains the relationship between two gray level in frequency [19], described by 8 features.
b) GLRLM: GLRLM examines runs of similar gray values in an image [20–22], described by 13 features.
c) GLSZM: GLSZM reflects the homogeneity of small groups [18,19], described by 13 features.
d) NGTDM: NGTDM describes the spatial changes in intensity [23], described by 5 features.

4) Wavelet decomposition features: in order to figure out the characteristics of the tumor in different frequencies, low-pass and high-pass filters have been designed to decompose the original images into 8 decompositions, and intensity and texture features computation have been applied to each decomposition, respectively.

2.4 Feature selection and heat map

Of all 555 radiomics features, the portion of features which show significant correlation with the age groups was expected to be found. So T test is applied to all features under the assumption that age divided according to the threshold of 45 years old.

After the remarkable features have been gotten, a heat map has been drawn using unsupervised hierarchical clustering of those features for the effective evaluation of age classification.

3. Results

3.1 Tumor segmentation

The segmentation of 96 T1 Contrast MRI glioma images have been done with the fully automatic method mentioned above. The novel segmentation method shows high accuracy and precision for all glioma images. As an example, one of the segmentation result is shown in Figure 1.

3.2 Feature extraction and selection

After the segmentation of tumors has been finished, 555 features in 96 tumors images have been extracted, which shows the characteristics of intensity, shape and texture of the tumors. In order to obtain the value of radiomics features to capture the age differences of glioblastoma, the T test has been applied to all features. 101 of 555 features have been found to perform correlation with age classification ($p$ value < .05), one of the most distinctive features’ distribution corresponding to the age range is shown in Figure 2. The top five highly correlated features are listed in Table 3.

It can be seen from Table 3 that features which show the most significant correlation to age division are the ones related to gray-level distribution and high order heterogeneity. It should be mentioned that the most discriminative feature in our experiment is kurtosis of decomposition HLH, accordant with another related study in lung cancer [12]. The result indicated...
that image feature after decomposing in HLH may exhibit more clinical value.

With the selected 101 features, we performed an unsupervised clustering showing the clusters of patients with similar radiomics features, as shown in Figure 3. Two main clusters of patients with age group division shown in Figure 3(c) have been compared, and remarkable correspondence \((p\ value = .006)\) has been found.

It could be seen from the clusters of the features as shown in Figure 3(b) that most of features selected are from wavelet features which express the performance of tumor heterogeneity in high order and different frequencies, in contrast, almost none of shape features show the difference between two age groups.

4. Discussion

Most previous studies on the relationship between ages and medical images are described with the qualitative analysis. In this study, the relationship between ages and medical images of glioblastoma patients has been explored in a quantitative and statistical way, which should be more convincing and conclusive. The use of quantitative image features in T1 Contrast MRI images has been demonstrated and significant correlations with age groups of glioblastoma patients has been reported. Results also draw the conclusion that radiomics features have the value of prognosis because it has been proved that age is an important factor of prognosis \([5–7]\).

Among the features with discrimination ability, a number of features reflected gray level distributions and high order heterogeneity of wavelet decomposition which are not visible to human eyes. The results are in accordance with the related researches’ results that gray-level distribution and intra-tumor heterogeneity exhibit the association with molecular level and tumor cell proliferation \([12,24]\), and proliferation, as a general hallmark of cancer, has been proved to be connected with age \([25]\).

Actually, it was widely recognized that glioblastoma patients who are younger live much more longer. Recent researches of glioblastoma have found that totally different genetic background was present in younger and elder population \([26]\). It may provide an interesting explanation to different tumor growth patterns in distinct age groups resulting from varied MRI features. Previous study has already demonstrated that isocitrate dehydrogenase 1 (IDH1) mutated glioblastomas were preferentially localized in left frontal lobe characterized with sharp tumor border \([27]\). It may be the first paper of our study to report MRI features by age stratification in glioblastoma. Our findings confirmed that glioblastoma with different age groups are genetically different brain tumors which should not be treated in the same way. However, the deficiency of our study is lack of survival data and its correlation with other clinical subtypes was not analyzed.

Indeed, it is often hypothesized that glioblastoma patients at different ages have different origins of the tumors \([8]\). The diseases of younger patients are suspected to evolve from low level gliomas such as

**Figure 2.** Boxplot result of a remarkable feature. \((p\ value = 0.001)\).

**Table 3.** Five highly correlated features to age group.

| Feature | \(p\) value |
|---------|-------------|
| Kurtosis of decomposition HLH (feature 378) | .001 |
| Minimum of intensity (feature 8) | .003 |
| Kurtosis of decomposition LLL (feature 138) | .003 |
| Histogram skewness of decomposition LHH (feature 448) | .003 |
| Long run high gray-level emphasis of decomposition of HHL (feature 295) | .004 |
astrocytoma, and on the contrary, elder patients’ diseases seem to be idiopathic.

In our study, the relationship between quantitative image features and patients’ age has been demonstrated, which may inspire the relevant researchers to look into the origins and growth patterns of glioblastoma at different ages in a non-invasive approach. There were some limitations in our study. Firstly, all the features were extracted from the whole regions of glioblastoma. However, the result might be promoted by dividing the tumor regions into more detailed classes such as edema, necrosis and other tumor regions. With more detailed segmentation results, the characteristics of the tumor regions could be better reflected. Another limitation in our study is the lack of glioblastoma cases. The analysis could be carried out more detailed with more data.

Further research could be progressed by following up the 96 glioblastoma patients to get the survival time of each patients and Kaplan-Meier survival analysis may be adopted to explore the possible association of radiomics features with survival, which may confirm the prognostic value of those quantitative features that we extracted and selected. It could provide a viable way to get some insight into the importance of individual features. On the other hand, radiomics is a potential technique to study the origin differences among glioblastomas with different ages at diagnosis, which will be the future work of our group.

5. Conclusion

In this work, the prognostic value of radiomics data by finding the relationship between ages and medical images has been explored, as it has been widely accepted that age is an important prognostic factor in clinical data. In order to get standardized and precise segmentation results of tumors, a novel fully
automatic segmentation method has been applied to 96 glioblastoma patients’ T1 Contrast MRI images. Then 555 radiomics features have been extracted from the tumor regions of medical images. With the purpose of investigating the value of radiomics features to capture the age difference, we selected 101 features which show consistent significance with age group (<45 and ≥45 years old) (p value <.05). So as to reveal the classification results of those radiomics features, an unsupervised clustering has been utilized to draw a heat map, eventually it is successful to find out the cluster results is corresponding with age group (p value = .006). In conclusion, glioblastoma in different age groups present different radiomics-feature patterns with statistical significance. Since radiomics features are highly associated with pathology, biomarkers and genomics, results in this paper indicates that glioblastoma in different age groups should have different pathologic, protein, or genic origin.

Acknowledgements
This work was supported by the National Basic Research Program of China (2015CB755500), the National Natural Science Foundation of China (61271071). Thanks to Dr. Zhifeng Shi, Dr. Liang Chen and Dr. Ying Mao for providing clinical guidance.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This work was supported by the National Basic Research Program of China (2015CB755500), the National Natural Science Foundation of China (61271071).

References
[1] Furnari FB, Fenton T, Bachoo RM, et al. Malignant astrocytic glioma: genetics, biology, and paths to treatment. Genes Dev. 2007;21:2683–2710.
[2] Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol. 2007;114:97–109.
[3] Sanai N, Berger MS. Glioma extent of resection and its impact on patient outcome. Neurosurgery. 2008;64:753–764.
[4] Buckner JC. Factors influencing survival in high-grade gliomas. Semin Oncol. 2003;30:10–14.
[5] Yang P, Wang YZ, Peng XX, et al. Management and survival rates in patients with glioma in China (2004–2010): a retrospective study from a single-institution. J Neurooncol. 2013;113:259–266.
[6] Li SW, Qiu XG, Chen BS, et al. Prognostic factors influencing clinical outcomes of glioblastoma multiforme. Chinese Med J-Peking. 2009;122:1245–1249.
[7] Catalan-Uribarrena G, Bilbao-Barandica G, Pomposo-Gaztelu I, et al. Prognostic factors and survival in a prospective cohort of patients with high-grade glioma treated with carmustine wafers or temozolomide on an intention-to-treat basis. Acta Neurochir. 2012;52:211–222.
[8] Korshunov A, Ryzhova M, Hovestadt V, et al. Integrated analysis of pediatric glioblastoma reveals a subset of biologically favorable tumors with associated molecular prognostic markers. Acta Neuropathol. 2015;129:1–10.
[9] Wang Y, Liu S, Fan X, et al. Age-associated brain regions in gliomas: a volumetric analysis. J Neurooncol. 2015;123:299–306.
[10] Seidel C, Dorner N, Osswald M, et al. Does age matter? - A MRI study on peritumoral edema in newly diagnosed primary glioblastoma. BMC Cancer. 2011;11:127.
[11] Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. Magn Reson Imaging. 2012;30:1234–1248.
[12] Aerts HJWL Velazquez ER, Leijenaar RTH, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006.
[13] Gutman DA, Cooper LAD, Hwang SN, et al. MR imaging predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set . Radiology. 2013;267:560–569.
[14] Bourgier C, Colinge T, Ailleres N, et al. Radiomics: definition and clinical development. Cancer Radiother. 2015;19:6–7.
[15] 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. 2011;95:190–198.
[16] Dice LR. Measures of the amount of ecologic association between species. Ecology. 1945;26:297–302.
[17] Vallières M, Freeman CR, Skamene S, et al. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys Med Biol. 2015;60:5471–5496.
[18] Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. Magn Reson Imaging. 2004;22:81–91.
[19] Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst, Man, Cybern. 1973;SMC3:610–621.
[20] Bocchino C, Carabellase A, Caruso T, et al. Use of gray value distribution of run lengths for texture analysis. Pattern Recogn Lett. 1990;11:415–419.
[21] Galloway MM. Texture analysis using gray level run lengths. Comput Graph Image Process. 1975;4:172–179.
[22] Thibault G, Fertil B, Navarro C, et al. Shape and texture indexes and gray level size zone matrix
application to cell nuclei classification. Int J Pattern Recogn. 2013;27:155–167.

[23] Amadasun M, King R. Textural features corresponding to textural properties. IEEE Trans Syst, Man, Cybern. 1989;19:1264–1274.

[24] Segal E, Sirlin CB, Ooi C, et al. Decoding global gene expression programs in liver cancer by noninvasive imaging. Nat Biotechnol. 2007;25:675–680.

[25] Oestreicher N, White E, Malone KE. Hormonal factors and breast tumor proliferation: do factors that affect cancer risk also affect tumor growth? Breast Cancer Res Treat. 2004;85:133–142.

[26] Li R, Li HL, Wei Y, et al. Genetic and clinical characteristics of primary and secondary glioblastoma is associated with differential molecular subtype distribution. Oncotarget. 2015;6:7318–7324.

[27] Wang Y, Zhang T, Li S, et al. Anatomical localization of isocitrate dehydrogenase 1 mutation: a voxel-based radiographic study of 146 low-grade gliomas. Eur J Neurol. 2014;22:348–354.