

Better efficacy in differentiating WHO grade II from III oligodendrogliomas with machine-learning than radiologist's reading from conventional T1 contrast-enhanced and fluid attenuated inversion recovery images

Sha-Sha Zhao

Xi'an Tangdu Hospital of No4 Military Medical University

Xiu-Long Feng

Xi'an Tangdu Hospital of No4 Military Medical University

Yu-Chuan Hu

Xi'an Tangdu Hospital of No4 Military Medical University

Yu Han

Xi'an Tangdu Hospital of No4 Military Medical University

Qiang Tian

Xi'an Tangdu Hospital of No4 Military Medical University

Ying-Zhi Sun

Xi'an Tangdu Hospital of No4 Military Medical University

Jie Zhang

Xi'an Tangdu Hospital of No4 Military Medical University

Xiang-Wei Ge

Student Brigade, Air Force Medical University

Si-Chao Cheng

Student Brigade, Air Force Medical University

Xiu-Li Li

Deepwise AI Lab, Deepwise Inc.

Li Mao

Deepwise AI Lab, Deepwise Inc.

Shu-Ning Shen

Department of Stomatology, PLA 984 Hospital

Lin-Feng Yan

Xi'an Tangdu Hospital of No4 Military Medical University

Guang-Bin Cui

Xi'an Tangdu Hospital of No4 Military Medical University

Wen Wang (✉ wangwen@fmmu.edu.cn)

Tangdu Hospital, Air Force Medical University <https://orcid.org/0000-0001-6473-4888>

Research article

Keywords: Oligodendrogliomas; Machine learning; Radiomics; Random forest (RF); Magnetic resonance imaging (MRI)

Posted Date: January 14th, 2020

DOI: <https://doi.org/10.21203/rs.2.9727/v5>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at BMC Neurology on February 7th, 2020. See the published version at <https://doi.org/10.1186/s12883-020-1613-y>.

Abstract

Background: The medical imaging to differentiate World Health Organization (WHO) grade II (ODG2) from III (ODG3) oligodendrogliomas still remains a challenge. We investigated whether combination of machine learning with radiomics from conventional T1 contrast-enhanced (T1CE) and fluid attenuated inversion recovery (FLAIR) magnetic resonance imaging (MRI) offered superior efficacy. **Methods:** Thirty-six patients with histologically confirmed ODGs underwent T1CE and 33 of them underwent FLAIR MR examination before any intervention from January 2015 to July 2017 were retrospectively recruited in the current study. The volume of interest (VOI) covering the whole tumor enhancement were manually drawn on the T1CE and FLAIR slice by slice using ITK-SNAP and a total of 1072 features were extracted from the VOI using 3-D slicer software. Random forest (RF) algorithm was applied to differentiate ODG2 from ODG3 and the efficacy was tested with 5-fold cross validation. The diagnostic efficacy of radiomics-based machine learning and radiologist's assessment were also compared. **Results:** Nineteen ODG2 and 17 ODG3 were included in this study and ODG3 tended to present with prominent necrosis and nodular/ring-like enhancement ($P < 0.05$). The AUC, ACC, sensitivity, and specificity of radiomics were 0.798, 0.735, 0.672, 0.789 for T1CE, 0.774, 0.689, 0.700, 0.683 for FLAIR, as well as 0.861, 0.781, 0.778, 0.783 for the combination, respectively. The AUCs of radiologists 1, 2 and 3 were 0.700, 0.687, and 0.714, respectively. The efficacy of machine learning based on radiomics was superior to the radiologists' assessment. **Conclusions:** Machine-learning based on radiomics of T1CE and FLAIR offered superior efficacy to that of radiologists in differentiating ODG2 from ODG3.

Background

Oligodendrogliomas (ODGs), predominantly occur in adults with a peak between 40 and 60 years of age, constitute 5-20% of all gliomas [1]. Patients with low-grade (ODG2) are slightly younger than those with high-grade, anaplastic tumors (ODG3) [2]. The co-deletion of the short arm of chromosome 1 (1p) and the long arm of chromosome 19 (19q) [3] occurs in about 60–90% of ODGs, thus making it the molecular hallmark for ODGs [1].

Calcification [4, 5] and the cortical-subcortical location [5, 6], most commonly in the frontal lobe [4], are regarded as the characteristic features of ODGs. In contrast to other low-grade gliomas (LGG), minimal to moderate enhancement and moderately increased perfusion are commonly seen in ODGs, making the differentiation of ODG2 from ODG3 difficult. Besides, ODG3 often shares the imaging features with ODG2 on conventional MRI, leading to unreliable tumor grade prediction. Edema, haemorrhage, cystic degeneration and contrast enhancement are more commonly seen in ODG3, but may also be seen in ODG2 [4]. Thus, a new medical imaging diagnostic strategy for differentiation of ODG2 from ODG3 needs to be developed.

Advanced imaging techniques, including DWI, perfusion imaging, MR spectroscopy and PET, are employed to obtain more sensitive diagnostic markers, however with unsatisfying efficacy. Diffusion restriction is seldom observed in ODG2 [6]. Averaged ADC values are reported to be lower in high grade glioma (HGG) than in LGG, however, ADC values of ODG3 are overlapped with that of ODG2, making DWI unreliable maker to distinguish them [7]. Using the cut-off value of 1.75 for relative cerebral blood volume (rCBV) ratio, HGG can be differentiated from LGG with a sensitivity of 95% [8]. Unfortunately, these findings may not be suitable for differentiating ODGs, because markedly elevated rCBV can also be observed in ODG2, thus, a reliable distinction can't be easily achieved [7, 9, 10]. This is due to the presence of the short capillary segments in ODGs [5] which may contribute to the relatively low specificity (70%) reported by Law et al [8]. Therefore, focally elevated rCBV does not necessarily

indicate ODG3. Besides, correlation of K^{trans} with tumor grade is even poorer than that of rCBV, and it is more commonly used to assess the treatment effects [11]. Taking together, the efficacies of advanced MRI techniques in differentiating ODG2 from ODG3 are limited.

Combining quantitative image features extracted from conventional T1-weighted contrast-enhanced (T1CE) and fluid attenuated inversion recovery (FLAIR) images with machine learning algorithms, radiomics can provide comprehensive information that is difficult to perceive with visual inspection [12, 13] and is commonly used in tumor diagnosis, staging and prognosis of tumors [14-20]. However, most previous studies were mainly focused on advanced MR techniques, the varied post-processing models, varied interpretation and evaluation criteria restricted their clinical applications. Except for their limited diagnostic powers, these advanced MRI techniques are not commonly available in some rural areas. However, the T1CE and FLAIR are widely-used in almost all hospitals as the image routine sequences for glioma diagnosis and staging. It is thus feasible to combine radiomics with T1CE and FLAIR to establish a practical and economical imaging solution for differentiating ODG2 from ODG3.

In this study, we aimed to evaluate the diagnostic power of machine-learning based on T1CE and FLAIR imaging radiomics in comparison with the radiologists' performance in differentiating ODG2 from ODG3.

Methods

Patients

This study was approved by our institutional review board and the requirement for informed consent was waived based on its retrospective nature. From January 2015 to July 2017, patients with confirmed ODGs were retrospectively and consecutively recruited. Tumors were classified according to 2007 WHO classification or 2016 WHO guidelines when enough information was available. The including criteria were, 1. patients underwent preoperative conventional MRI scan. 2. patients underwent gross total or subtotal tumor resection and a confirmative pathological diagnosis was made. Thirty-six patients with TICE were included (19 men, 17 women; mean age =45 years; age range =9 - 65 years) and classified into two groups: ODG2 (n = 19; mean age = 46 years, age range =10 - 65 years) and ODG3 (n = 17; mean age = 44 years, age range = 9 - 65 years). Thirty-three out of the above 36 patients with FLAIR were enrolled (18 men, 15 women; mean age = 45 years; age range = 9 - 65 years) and classified into two groups: ODG2 (n = 17; mean age = 45 years, age range =10 - 65 years) and ODG3 (n = 16; mean age = 45 years, age range = 9 - 65 years). The patient selection is summarized in **Figure 1**.

MRI Data Acquisition

All patients underwent 3-T MR scanning (Discovery MR750, General Electric Medical System, Milwaukee, WI, USA) with an 8-channel head coil (General Electric Medical System). The initial routine scan sequences for each patient included T1-weighted imaging (T1WI) performed before and after contrast enhancement, an axial T2-weighted imaging (T2WI), and a transverse FLAIR to assist with diagnosis.

The parameters of the conventional MRI sequences were as the follows: T1WI with gradient echo (TR/TE, 1750 ms/24 ms; matrix size, 256 × 256; FOV, 24 × 24 cm; number of excitation, 1; slice thickness, 5 mm; gap, 1.5 mm), T2WI with turbo spin-echo (TR/TE, 4247 ms/93 ms; matrix size, 512 × 512; FOV, 24 × 24 cm; number of excitation, 1; slice thickness, 5 mm; gap, 1.5 mm) and sagittal T2WI (TR/TE, 10,639 ms/96 ms; matrix size, 384 × 384; FOV, 24 × 24 cm; number of excitation, 2; slice thickness, 5 mm; gap, 1.0 mm). We obtained axial FLAIR with the

following parameters: TR/TE, 8000 ms/165 ms; matrix size, 256 × 256; FOV, 24 × 24 cm; number of excitations, 1; slice thickness, 5 mm; gap, 1.5 mm.

Finally, T1CE were performed after intravenous bolus injection of gadodiamide (Omniscan; GE Healthcare, Co. Cork, Ireland), at a dose of 0.1 mmol/kg body weight. The parameters of T1CE with volumetric interpolated breath-hold examination (VIBE) were as the follows: TR/TE, 8.2 ms/3.2 ms; T1, 450 ms; flip angle 12°; section thickness, 1.2 mm; FOV, 24 × 24 cm; matrix size, 256 × 256; number of excitations, 1; image number, 140.

Tumor Segmentation or Delineation

Two neuroradiologists (S.S.Z with 8 years of experience and L.F.Y, with 12 years of experience in neuro-oncology imaging) independently reviewed all images. A third senior neuroradiologist (G.B.C, with 25 years of experience in neuro-oncology imaging) re-examined the images and determined the final imaging diagnoses when inconsistency occurred. The preoperative conventional image features of tumor were retrieved based on the criteria outlined in **Supplementary Table 1** (*online*).

The volumes of interest (VOIs) were semi-automatically segmented using ITK-SNAP (version3.6, <http://www.itk-snap.org>) by two neuroradiologists (S.S. Z and L.F.Y). The VOIs covering the enhanced lesion were drawn slice by slice on T1CE and co-registered to and FLAIR images, avoiding the regions of macroscopic necrosis, cyst, edema and non-tumor macrovessels [21].

Radiomics Strategy

Feature extraction Texture features include 162 first-order logic features, 216 gray level co-occurrence matrix (GLCM) features, 144 gray level run length matrix (GLRLM) features, 144 gray level size zone matrix (GLSZM) features, 126 grey level difference matrix (GLDM) features, 45 neighborhood grey-tone difference matrix (NGTDM) features and 14 shape Features. A total of 1072 features were extracted from the T1CE and FLAIR images using 3D-slicer software. We used the aforementioned features because these features were found to be relevant for distinguishing ODG2 from ODG3 in our previous studies by using MR imaging [16].

Feature selection After being centered and scaled, the highly redundant and correlated features were subjected to a two-step feature selection procedure. First, highly correlated features were eliminated using Pearson correlation analysis, with the r threshold of 0.75. Then, a random forest (RF) classifier consisting of a number of decision trees was used to rank the feature importance. Every node in the decision trees is a condition on a single feature, designed to split the dataset into two so that similar response values end up in the same set. The measurement based on which optimal condition is chosen is called impurity. For classification, it is typically either Gini impurity or information gain/entropy. Thus, when training a tree, it can be computed how much each feature decreases the weighted impurity in a tree. To build the RF, the impurity decrease from each feature can be averaged and the features are ranked according to this measurement. In our study, Gini impurity decrease was used as the criterion to indicate the feature importance.

Radiomics model building The 30 most important features were fed into a Conditional Inference RF classifier to build model [22]. Five-fold cross validation was employed for tuning hyperparameter number of RF trees. Five-fold cross validation including pre-processing, feature selection and model construction were performed 3 times in order to avoid bias and overfitting as much as possible. The final results were the average from 3 performances. There was no feature selection in the combination of T1CE and FLAIR throughout the model building. Accuracy,

sensitivity and specificity were computed to evaluate the classifying performance. The receiver operating characteristic (ROC) curve was also built to provide the area under the ROC curve (AUC). The larger the AUC, the better the classification [23]. The whole procedure of feature extraction and machine learning was described in **Figure 2**.

Radiologist's assessment To compare the efficacies of neuroradiologist and machine learning in differentiating ODG2 from ODG3, the images were also independently assessed by three junior neuroradiologists (X.L.F, G.X and Y.H with 6, 7 and 7 years of neuroradiology experience, respectively). The neuroradiologists were blinded to the clinical information, but were aware that the tumors were either ODG2 or ODG3, without knowing the exact number of patients with each entity. The three readers assessed only conventional MR images (T1WI, T2WI, FLAIR and T1CE), and recorded the final diagnosis using a 4-point scale (1 = definite ODG2; 2 = likely ODG2; 3 = likely ODG3; and 4 = definite ODG3) [24].

Statistical Analysis

Fisher exact test or the Chi-square test were used for the categorical variables and unpaired Student *t* test was used for continuous variable between ODG2 and ODG3 groups. The statistical analyses of clinical characteristics were performed by using SPSS 20.0 software (SPSS Inc., Chicago, IL, USA).

The statistical analyses of machine-learning were performed using R version 3.4.2 (R Foundation for Statistical Computing). A RF analysis was performed to train the machine-learning classifier. The goal of machine learning was to build the model to differentiate ODG2 from ODG3 based on radiomics features of T1CE and FLAIR images. The following R packages were used: the random forest package was used for feature ranking; the caret and unbalanced packages were used for RF classification. Classifier performance was determined by using accuracy, sensitivity and specificity. The AUC values were also calculated for three readers and compared with that of the radiomics classifier. *P* value < 0.05 was considered as statistical significance.

Results

Patient Characteristics

The main clinical characteristics and conventional MRI features of the 36 patients (ODG2 and ODG3) were summarized in **Table 1**. Tumor necrosis was more frequent in ODG3 than in ODG2 groups (*P* = 0.044), reflecting the hypoxia as a result of the rapid tumor growth. In addition, ODG3 were related to the nodular/ring-like enhancement patterns (*P* = 0.002). Besides, 10/19 (52.6%) of ODG2 and 10/17 (58.8%) of ODG3 situated in the frontal lobe, indicating no significant group difference. No significant difference of other clinical characteristics (gender, age) or imaging paradigms was observed between ODG2 and ODG3 patients.

Quantitative MR Histogram and Texture Features Analysis

The relative importance of features computed by using the Gini index to differentiate ODG2 from ODG3 was depicted in **Figure 3**. It can be seen that if all the high-throughput features were put into the RF classifiers, the classification performance could not be significantly improved because of the feature redundancy.

The strong relationship between radiomics features to differentiate ODG2 from ODG3 was also indicated in the radiomics heat map (**Figure 4**). The RF based feature selection strategy improved the performance of RF classifier.

After RF feature selection, 30 optimal features were selected to differentiate ODG2 from ODG3, with comparable efficacy to that of using all features.

Evaluation of Principal Components

When ODG2 and ODG3 were differentiated by using principal components, similar tumor tissue formed characteristic clusters. These clusters, although heterogeneous, defined a specific VOI (eg, **Figure 5**) and were separable from other tumors (clusters). More important, the calculated principal components of the VOIs from ODG2 and ODG3 allowed clear separation of these two important regions.

Diagnostic performance of Radiomics and Radiologists

The performance of radiomics and 3 radiologists in differentiating ODG2 from ODG3 was also compared. **Table 2** and **Figure 6** summarized the diagnostic performance of the radiomic features derived by using MR images from T1CE, FLAIR and their combination to distinguish ODG2 from ODG3. Radiomic features from their combination showed significantly better diagnostic performance than that of FLAIR or T1CE. Violin plots graphed for the first 9 radiomic features derived from T1CE, FLAIR and their combination were presented in **Figure 6**. The AUC, sensitivity, specificity and accuracy of radiomics were 0.798 (95%CI 0.699-0.896), 0.672, 0.789, 0.735 for T1CE, 0.774 (95%CI 0.671-0.877), 0.700, 0.683, 0.689 for FLAIR, and 0.861 (95%CI 0.783-0.940), 0.778, 0.783, 0.781 for their combination, respectively. The AUCs of the three radiologists were 0.700 (95%CI 0.519-0.880), 0.687 (95%CI 0.507-0.867) and 0.714 (95%CI 0.545-0.883) for readers 1, 2 and 3, respectively. The radiomics classifier performed superior to the 3 junior radiologists. The representative cases of ODG2 and ODG3 were presented in **Figure 7**. The clinical application of radiomics-based machine learning could be justified based on our findings.

Discussion

Radiomics is an emerging field that treats images as data rather than pictures and analyzes a large number of features extracted from 1 image in relation to clinical variables of interest. A few studies on radiomics analyses of glioma have been published over the last years and advocated for machine learning models in predicting tumor histology and grade [25]. Radiomics has been suggested as a robust strategy to noninvasively classify lesions [14, 26]. This work suggested that radiomics from T1CE and FLAIR can be useful for differentiating ODG2 from ODG3, with the superior efficacy to that of radiologists, thus, its clinical application could be justified based on the current study.

From the angle of experiment design, there are three aspects worthy noting in this study. First, the 'real world' data were used to test our scientific hypothesis. Second, all images analyzed in the current study were taken exclusively from routine clinical diagnostic scans. Third, based on the social-economic consideration, the levels of accuracy were based on the radiomics of commonly available T1CE and FLAIR images, without an acquisition of spectroscopy, CBV or perfusion information, all of which would prolong the scanning time and increase economic burden to patients. Upon our expectation, the radiomics strategy performed superior to that of radiologists.

The reasons for the improved diagnostic performance of radiomics are as the following. First, radiomic methods, given their ability to discern patterns and combine information in a way that humans cannot, showed substantial promise for the future of radiology and precision medicine [27]. However, radiologists distinguished ODG2 from ODG3 by visual diagnosis using rough information from T1CE and FLAIR. Second, it has been reported that the

performance of an SVM classifier can be significantly reduced by the inclusion of redundant features and this effect is more obvious for a small training set [28]. In this study, it was found that the combination of conventional T1CE and FLAIR features provided lower classification error than features of individual sequence, which may thus emphasize the importance of using a multiparametric approach. In addition, highly correlated features were eliminated using Pearson correlation analysis, which was also further ranked by using the random forest classifier consisting of a number of decision trees. This indicated that redundant features removed can have a contribution to the classification of ODG2 and ODG3.

Radiomics strategy not only performed superior to radiologists, but also could be used as an auxiliary means to overcome some problems attained to radiologists. First of all, the frequency of interruptions during a reporting session is associated with up to 13% increase in time for reporting and an increased potential for errors [29]. Then, fatigue adversely impacts the visual system including: worse accommodation, decreased saccadic velocity and reduced gaze volume and coverage [30]. At last, a number of cognitive biases may adversely affect the accuracy of a radiologists report of a glioma [31]. In order to reduce reporting time and cognitive biases, both of which may lead to reporting and diagnostic errors, radiomics offers a significant advantage [32], particularly in the context of general radiologists who may lack expertise in neuro-oncology. Nevertheless, the current radiomics strategy involves too much pre- and post-process before the suitable machine learning model is established, more studies focusing on the efficacy-cost balance of such a machine learning system should be further conducted before its clinical application.

Furthermore, a few limitations of this study should be noticed. In the first place, sample number of the patients is relatively small. Although current results of 5-fold cross validation showed that the evaluation of diagnostic efficacy were robust despite the relatively small sample size, which did not cause the classifier to be skewed towards a particular class. It is desirable to verify the classifier on a larger data size in the future. Besides, this radiomic method incorporated vessel removal in its methodology, this method may fail for certain cases that were non-tumor vessels intertwined with tumor vessels. Signal intensity curves of prominent vessels can be used as a differentiating feature for such cases. The last, a continuous effort on enlarging the dataset so as to test its external validation is required.

Conclusions

In conclusion, this study demonstrates our findings that use of a machine learning algorithm, derived from 'real word' T1CE and FLAIR images, which can differentiate ODG2 from ODG3 in newly diagnosed gliomas with a superior efficacy to that of radiologists. The RF selected features can reduce the labor in applying this strategy, and the strategy can be applied clinic based on our findings.

List Of Abbreviations

Abbreviation	Full name of English
ODG	Oligodendroglioma
HGG	High Grade Glioma
LGG	Low Grade Glioma

rCBV	Relative Cerebral Blood Volume
T1WI	T1-weighted imaging
T2WI	T2-weighted imaging
FLAIR	Fluid-Attenuated Inversion recovery
T1CE	T1-weighted contrast-enhanced image
RF	Random Forest
BBB	Brain Blood Barrier

Declarations

Ethics approval and consent to participate

This is a retrospective study that does not require the approval of the ethics committee. (Not applicable)

Consent to publication

Our manuscript does not contain any individual person's data. (Not applicable)

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study received financial support from the National key research and development program of China (No. 2016YFC0107105 to Dr. Cui G.B.), the Science and Technology Development of Shaanxi Province (No. 2014JZ2-007 to Dr. Cui G.B.; 2015kw-039 to Dr. Wang W) and Innovation and Development Foundation of Tangdu Hospital (No. 2016LCYJ001 to Dr. Cui G.B.) and Intramural Grant of Tangdu Hospital (Drs. Yan LF and Wang W). The funding body played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions

WW and CGB conceived the project, ZSS, YLF, FXL, CSC and HYC conducted the patient enrollment and data collection, HY, TQ, SYZ, ZJ, GXW, SSN, LXL and ML contributed to the data analysis and graph making, ML and LXL contributed to the thoughtful discussion and constructive help in data analysis. ZSS and WW drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank Drs Xue-Bin Lei, Sai Wang, Jin Zhang, Ying Yu, Qian Sun from Department of Radiology, Tangdu Hospital and Dr Xiao-Cheng Wei from GE healthcare for their great contribution to this work.

Authors' Information

ZSS MD YLF MD & Ph.D. FXL MD HYC MD HY MD
TQ MD SYZ MD ZJ MD GXW MD CSC MD CGB MD & Ph.D. WW MD &
Ph.D. SSN MD LXL Ph.D. ML BE

References

1. Van Den Bent MJ, Bromberg JE, Buckner J. Low-grade and anaplastic oligodendroglioma. *Handb Clin Neurol*. 2016; 134:361-380. <http://doi.org/10.1016/B978-0-12-802997-8.00022-0>.
2. Bromberg JE, van den Bent MJ. Oligodendrogliomas: molecular biology and treatment. *Oncologist*. 2009; 14(2):155-163. <http://doi.org/10.1634/theoncologist.2008-0248>.
3. Jenkins RB, Blair H, Ballman KV *et al*. A t(1;19)(q10;p10) mediates the combined deletions of 1p and 19q and predicts a better prognosis of patients with oligodendroglioma. *Cancer Res*. 2006; 66(20):9852-9861. <http://doi.org/10.1158/0008-5472.CAN-06-1796>.
4. Koeller KK, Rushing EJ. From the archives of the AFIP: Oligodendroglioma and its variants: radiologic-pathologic correlation. *Radiographics*. 2005; 25(6):1669-1688. <http://doi.org/10.1148/rg.256055137>.
5. Louis DN, Ohgaki H, Wiestler OD *et al*. The 2007 WHO classification of tumours of the central nervous system. *Acta neuropathologica*. 2007; 114(2):97-109. <http://doi.org/10.1007/s00401-007-0243-4>.
6. Osborn AG: Osborn's Brain: Imaging, Pathology, and Anatomy (1st edition). Salt Lake City, UT: Amirsys, Inc.; 2012.
7. Al-Okaili RN, Krejza J, Wang S, Woo JH, Melhem ER. Advanced MR imaging techniques in the diagnosis of intraaxial brain tumors in adults. *Radiographics*. 2006; 26 Suppl 1:S173-189. <http://doi.org/10.1148/rg.26si065513>.
8. Law M, Yang S, Wang H *et al*. Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. *AJNR Am J Neuroradiol*. 2003; 24(10):1989-1998. <http://doi.org>.
9. Lev MH, Ozsunar Y, Henson JW *et al*. Glial tumor grading and outcome prediction using dynamic spin-echo MR susceptibility mapping compared with conventional contrast-enhanced MR: confounding effect of elevated rCBV of oligodendrogliomas [corrected]. *AJNR Am J Neuroradiol*. 2004; 25(2):214-221. <http://doi.org>.
10. Chawla S, Wang S, Wolf RL *et al*. Arterial spin-labeling and MR spectroscopy in the differentiation of gliomas. *AJNR Am J Neuroradiol*. 2007; 28(9):1683-1689. <http://doi.org/10.3174/ajnr.A0673>.
11. Lacerda S, Law M. Magnetic resonance perfusion and permeability imaging in brain tumors. *Neuroimaging Clin N Am*. 2009; 19(4):527-557. <http://doi.org/10.1016/j.nic.2009.08.007>.
12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016; 278(2):563-577. <http://doi.org/10.1148/radiol.2015151169>.

13. Prasanna P, Patel J, Partovi S, Madabhushi A, Tiwari P. Radiomic features from the peritumoral brain parenchyma on treatment-naive multi-parametric MR imaging predict long versus short-term survival in glioblastoma multiforme: Preliminary findings. 2017; 27(10):4188-4197. <http://doi.org10.1007/s00330-016-4637-3>.
14. Huang YQ, Liang CH, He L *et al*. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. *J Clin Oncol*. 2016; 34(18):2157-2164. <http://doi.org10.1200/JCO.2015.65.9128>.
15. Horvat N, Veeraraghavan H, Khan M *et al*. MR Imaging of Rectal Cancer: Radiomics Analysis to Assess Treatment Response after Neoadjuvant Therapy. 2018; 287(3):833-843. <http://doi.org10.1148/radiol.2018172300>.
16. Tian Q, Yan LF, Zhang X. Radiomics strategy for glioma grading using texture features from multiparametric MRI. 2018. <http://doi.org10.1002/jmri.26010>.
17. Kalinli A, Sarikoc F, Akgun H, Ozturk F. Performance comparison of machine learning methods for prognosis of hormone receptor status in breast cancer tissue samples. *Comput Methods Programs Biomed*. 2013; 110(3):298-307. <http://doi.org10.1016/j.cmpb.2012.12.005>.
18. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJ. Machine Learning methods for Quantitative Radiomic Biomarkers. *Scientific reports*. 2015; 5:13087. <http://doi.org10.1038/srep13087>.
19. Chae HD, Park CM, Park SJ, Lee SM, Kim KG, Goo JM. Computerized texture analysis of persistent part-solid ground-glass nodules: differentiation of preinvasive lesions from invasive pulmonary adenocarcinomas. *Radiology*. 2014; 273(1):285-293. <http://doi.org10.1148/radiol.14132187>.
20. Vamvakas A, Williams SC, Theodorou K *et al*. Imaging biomarker analysis of advanced multiparametric MRI for glioma grading. *Phys Med*. 2019; 60:188-198. <http://doi.org10.1016/j.ejmp.2019.03.014>.
21. Yushkevich PA, Yang G, Gerig G. ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. *Conf Proc IEEE Eng Med Biol Soc*. 2016; 2016:3342-3345. <http://doi.org10.1109/EMBC.2016.7591443>.
22. Tagliamonte SA, Baayen RH. Models, forests and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation & Change*. 2012; 24(2):135-178. <http://doi.org>.
23. Cui Z, Xia Z, Su M, Shu H, Gong G. Disrupted white matter connectivity underlying developmental dyslexia: A machine learning approach. *Hum Brain Mapp*. 2016; 37(4):1443-1458. <http://doi.org10.1002/hbm.23112>.
24. Suh HB, Choi YS, Bae S *et al*. Primary central nervous system lymphoma and atypical glioblastoma: Differentiation using radiomics approach. *Eur Radiol*. 2018; 28(9):3832-3839. <http://doi.org10.1007/s00330-018-5368-4>.
25. Takahashi S, Takahashi W, Tanaka S *et al*. Radiomics Analysis for Glioma Malignancy Evaluation Using Diffusion Kurtosis and Tensor Imaging. *Int J Radiat Oncol Biol Phys*. 2019. <http://doi.org10.1016/j.ijrobp.2019.07.011>.
26. Aerts HJ, Velazquez ER, Leijenaar RT *et al*. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014; 5:4006. <http://doi.org10.1038/ncomms5006>.
27. Rudie JD, Rauschecker AM, Bryan RN, Davatzikos C, Mohan S. Emerging Applications of Artificial Intelligence in Neuro-Oncology. *Radiology*. 2019; 290(3):607-618. <http://doi.org10.1148/radiol.2018181928>.
28. Sengupta A, Ramaniharan AK, Gupta RK, Agarwal S, Singh A. Glioma grading using a machine-learning framework based on optimized features obtained from T1 perfusion MRI and volumes of tumor components.

J Magn Reson Imaging. 2019; 50(4):1295-1306. <http://doi.org10.1002/jmri.26704>.

29. Williams LH, Drew T. Distraction in diagnostic radiology: How is search through volumetric medical images affected by interruptions? *Cogn Res Princ Implic*. 2017; 2(1):12. <http://doi.org10.1186/s41235-017-0050-y>.
30. Waite S, Kolla S, Jeudy J *et al*. Tired in the Reading Room: The Influence of Fatigue in Radiology. *J Am Coll Radiol*. 2017; 14(2):191-197. <http://doi.org10.1016/j.jacr.2016.10.009>.
31. Lee CS, Nagy PG, Weaver SJ, Newman-Toker DE. Cognitive and system factors contributing to diagnostic errors in radiology. *AJR Am J Roentgenol*. 2013; 201(3):611-617. <http://doi.org10.2214/AJR.12.10375>.
32. Thrall JH, Li X, Li Q *et al*. Artificial Intelligence and Machine Learning in Radiology: Opportunities, Challenges, Pitfalls, and Criteria for Success. *J Am Coll Radiol*. 2018; 15(3 Pt B):504-508. <http://doi.org10.1016/j.jacr.2017.12.026>.

Tables

Table 1 Clinical characteristics and MRI features of patients

Variable	ODG2	ODG3	Total	P value
No. of patients, n	19	17	36	NA
Location, n (%)				0.378
Frontal	10/19 (52.6)	10/17 (58.8)	20/36 (55.6)	
Temporal	3/19 (15.8)	5/17 (29.4)	8/36 (22.2)	
Parietal	3/19 (15.8)	1/17 (5.9)	4/36 (11.1)	
Insular	1/19 (5.3)	1/17 (5.9)	2/36 (5.6)	
Occipital	0/19 (0)	0/17 (0)	0/36 (0)	
Others	2/19 (10.5)	0/17 (0)	2/36 (5.6)	
Gender, n (%)				0.202
Male	8/19 (42.1)	11/17 (64.7)	19/36 (52.8)	
Female	11/19 (57.9)	6/17 (35.3)	17/36 (47.2)	
Age ^a				0.788
Mean \pm SD	45.6 \pm 13.7	44.3 \pm 15.1	45.0 \pm 14.4	
Signal, n (%)				0.092
Homogeneous	6/19 (31.6)	1/17 (5.9)	7/36 (19.4)	
Heterogeneous	13/19 (68.4)	16/17 (94.1)	29/36 (80.6)	
Tumor cross midline, n (%)				1.000
No	16/19 (84.2)	14/17 (82.4)	30/36 (83.3)	
Yes	3/19 (15.8)	3/17 (17.6)	6/36 (16.7)	
Multiple foci, n (%)				0.736
No	12/19 (63.2)	9/17 (52.9)	21/36 (58.3)	
Yes	7/19 (36.8)	8/17 (47.1)	15/36 (41.7)	
Necrosis, n (%)				0.044*
No	13/19 (68.4)	5/17 (29.4)	18/36 (50.0)	
Yes	6/19 (31.6)	12/17 (70.6)	18/36 (50.0)	
Cyst, n (%)				0.255
No	16/19 (84.2)	11/17 (64.7)	27/36 (75.0)	
Yes	3/19(15.8)	6/17 (35.3)	9/36 (25.0)	
Edema, n (%)				0.106
No	4/19 (21.1)	0/17 (0)	4/36(11.1)	
Yes	15/19 (78.9)	17/17(100.0)	32/36 (88.9)	
Border, n (%)				1.000
Sharp/smooth	2/19 (10.5)	1/17 (5.9)	3/36 (8.3)	
Indistinct/irregular	17/19 (89.5)	16/17 (94.1)	33/36 (91.7)	
Enhancement, n (%)				0.002*
No/blurry	15/19 (78.9)	4/17 (23.5)	19/36 (52.8)	
Nodular/ring-like	4/19 (21.1)	13/17 (76.5)	17/36 (47.2)	
Cognitive dysfunction, n (%)				0.274
No	7/19 (36.8)	3/17 (17.6)	10/36 (27.8)	
Yes	12/19 (63.2)	14/17 (82.4)	26/36 (72.2)	
Epileptic seizures, n (%)				1.000
No	10/19 (52.6)	9/17 (52.9)	19/36 (52.8)	
Yes	9/19 (47.4)	8/17 (47.1)	17/36 (47.2)	

Table 2 Statistical differences between oligodendrogliomas (ODG2 and ODG3) between radiomic features determined RF classifier

Feature	Gini Importance	ODG2		ODG3		<i>P</i> Value
		Median	Interquartile Range	Median	Interquartile Range	
Long Run High Grey Level Emphasis _All Direction_offset5_SD	1.66	94.01	75.17-99.68	62.77	49.68-69.65	0.020*
Correlation _All Direction_offset9_SD	1.52	11.4 (10 ⁻⁵)	6.11-12.58 (10 ⁻⁵)	6.13 (10 ⁻⁵)	3.05-9.54 (10 ⁻⁵)	0.011*
Long Run Low Grey Level Emphasis _All Direction_offset7_SD	1.06	2.28 (10 ⁻⁵)	0.18-1.06 (10 ⁻⁵)	7.36 (10 ⁻⁵)	0.51-3.96 (10 ⁻⁵)	0.339
Short Run High Grey Level Emphasis _All Direction_offset6_SD	0.94	23.26	17.65-29.72	14.42	5.94-21.34	0.016*
RMS	0.92	930.73	818.53-1037.73	1213.40	949.17-1493.34	0.014*
Long Run High Grey Level Emphasis _All Direction_offset7_SD	0.90	469.55	285.36-601.49	285.89	178.74-362.53	0.009*
GLCM Energy _All Direction_offset4_SD	0.62	2.80 (10 ⁻⁵)	0.87-2.78 (10 ⁻⁵)	1.63 (10 ⁻⁵)	1.14-2.02 (10 ⁻⁵)	0.223
Short Run Low Grey Level Emphasis _All Direction_offset3_SD	0.59	0.37 (10 ⁻⁵)	0.01-0.08 (10 ⁻⁵)	0.96 (10 ⁻⁵)	0.62-0.76 (10 ⁻⁵)	0.231
GLCM Entropy _All Direction_offset5_SD	0.58	0.32	0.09-0.38	0.18	0.11-0.20	0.134
Compactness2	0.52	16.16	15.23-17.10	19.77	15.86-26.80	0.020*
Haralick Correlation _angle135_offset9	0.52	5470.3 (10 ⁵)	2510.0-7750.0 (10 ⁵)	4880.59 (10 ⁵)	2305.0-4335.0 (10 ⁵)	0.692
Run Length Nonuniformity _All Direction_offset3_SD	0.49	125.50	42.61-117.46	180.04	20.69-327.44	0.383
Intensity Variability	0.45	11.45	7.98-14.75	9.04	2.75-11.69	0.182
Inverse Difference Moment _All Direction_offset9_SD	0.44	570.91 (10 ⁻⁵)	338.84-731.06 (10 ⁻⁵)	495.23 (10 ⁻⁵)	337.92-600.48 (10 ⁻⁵)	0.449
Age	0.44	46	43-55	44	37-57	0.788
Long Run Low Grey Level Emphasis _All Direction_offset1_SD	0.43	2.12 (10 ⁻⁵)	0.22-1.80 (10 ⁻⁵)	8.62 (10 ⁻⁵)	0.16-4.39 (10 ⁻⁵)	0.177
High Grey Level Run Emphasis _All Direction_offset9_SD	0.42	4.26	2.64-5.14	6.77	3.02-9.50	0.099
High Grey Level Run Emphasis _All Direction_offset6_SD	0.41	6.02	2.55-7.28	7.18	3.80-7.63	0.486
kurtosis	0.40	2.22	0.05-1.37	0.69	-0.09-1.51	0.359
GLCM Entropy _All Direction_offset7_SD	0.40	0.24	0.14-0.33	0.37	0.09-0.25	0.512
GLCM Entropy_angle135_offset9	0.39	7.69	6.66-8.70	8.18	6.27-10.19	0.464
Short Run Emphasis _All Direction_offset8_SD	0.39	130.87 (10 ⁻⁵)	63.37-175.07 (10 ⁻⁵)	110.70 (10 ⁻⁵)	66.12-147.02 (10 ⁻⁵)	0.423
Minimum Intensity	0.38	114.79	17.0-145.0	149.82	7.0-275.0	0.558
Correlation_angle135_offset9	0.38	43.92 (10 ⁻⁵)	29.65-59.18 (10 ⁻⁵)	35.82 (10 ⁻⁵)	20.61-51.89 (10 ⁻⁵)	0.217
High Grey Level Run Emphasis _All Direction_offset8_SD	0.36	5.42	2.48-7.41	7.21	2.74-11.59	0.354
Short Run Low Grey Level Emphasis _All Direction_offset1_SD	0.35	1.51 (10 ⁻⁵)	0.08-1.31 (10 ⁻⁵)	2.98 (10 ⁻⁵)	0.10-3.16 (10 ⁻⁵)	0.316
GLCM Entropy_angle90_offset8	0.34	8.27	7.19-9.14	8.11	6.46-9.73	0.794
Cluster Shade_angle135_offset3	0.29	28109.70	-47175.30-65223.50	250959.5	-23039.0-229078.0	0.180
Cluster Shade _All Direction_offset7_SD	0.29	27784.98	4378.85-26851.50	29946.8	6717.39-22431.35	0.920
GLCM Entropy _All Direction_offset9_SD	0.26	0.31	0.18-0.37	0.22	0.13-0.35	0.110

Note: Feature relevance was assessed by using mean decrease in Gini index-based feature importance averaged over 100 trials. *P* values are adjusted for false-discovery rate by using Benjamini-Hochberg method. ODG2 = oligodendroglioma, ODG3 = anaplastic oligodendroglioma, RF = random forest.

Table 3 Diagnostic performance of comparison of radiomics and human assessment

	Sensitivity	Specificity	AUC	ACC
Radiomics (T1CE)	0.672	0.789	0.798 (95 % CI: 0.699, 0.896)	0.735
Radiomics (FLAIR)	0.700	0.683	0.774 (95 % CI: 0.671, 0.877)	0.689
Radiomics (T1CE+FLAIR)	0.778	0.783	0.861 (95 % CI: 0.783, 0.940)	0.781
Reader1	0.824	0.632	0.700 (95 % CI: 0.519, 0.880)	0.722
Reader2	0.706	0.684	0.687 (95 % CI: 0.507, 0.867)	0.694
Reader3	0.647	0.632	0.714 (95 % CI 0.545-0.883)	0.667

Figures

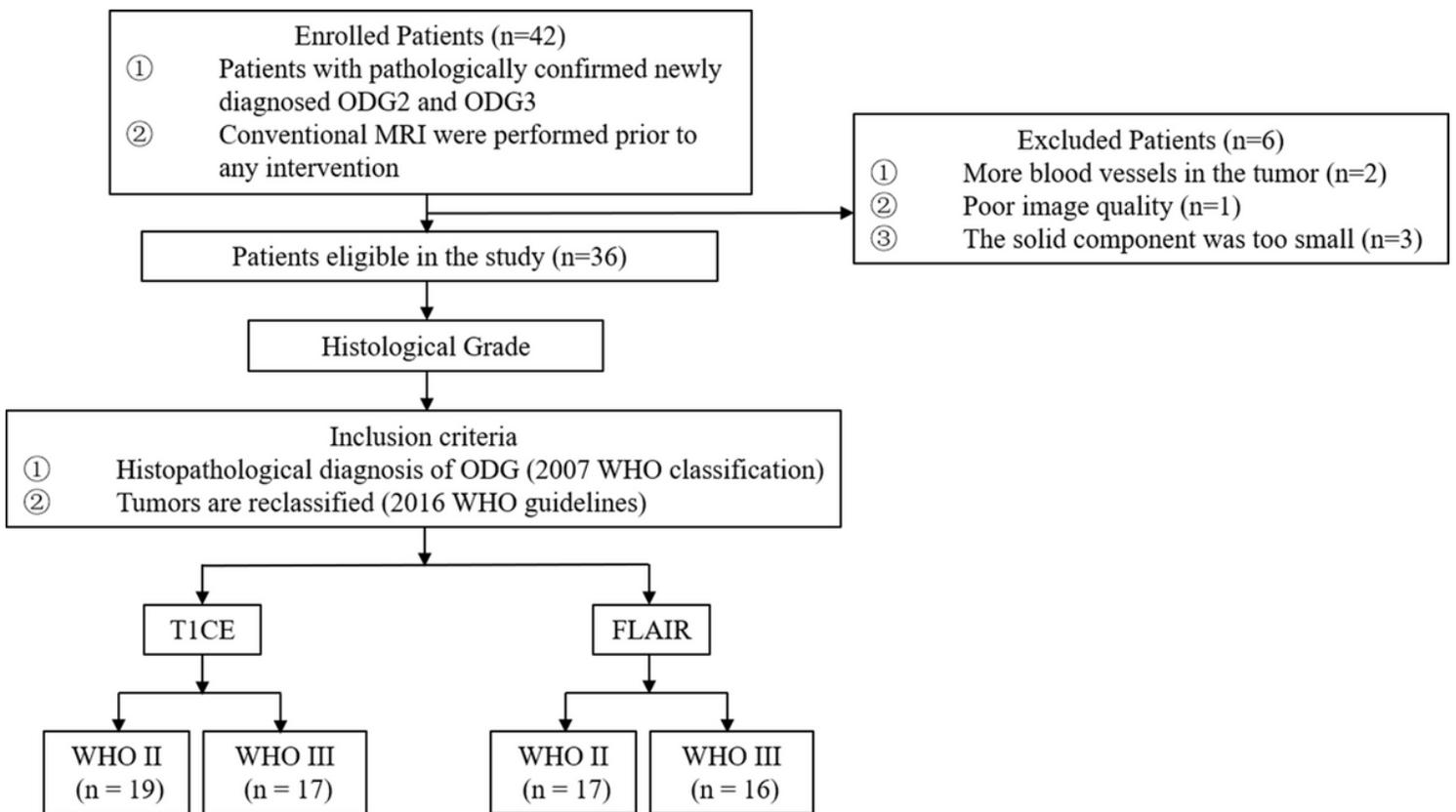


Figure 1

Flow diagram of the study design.

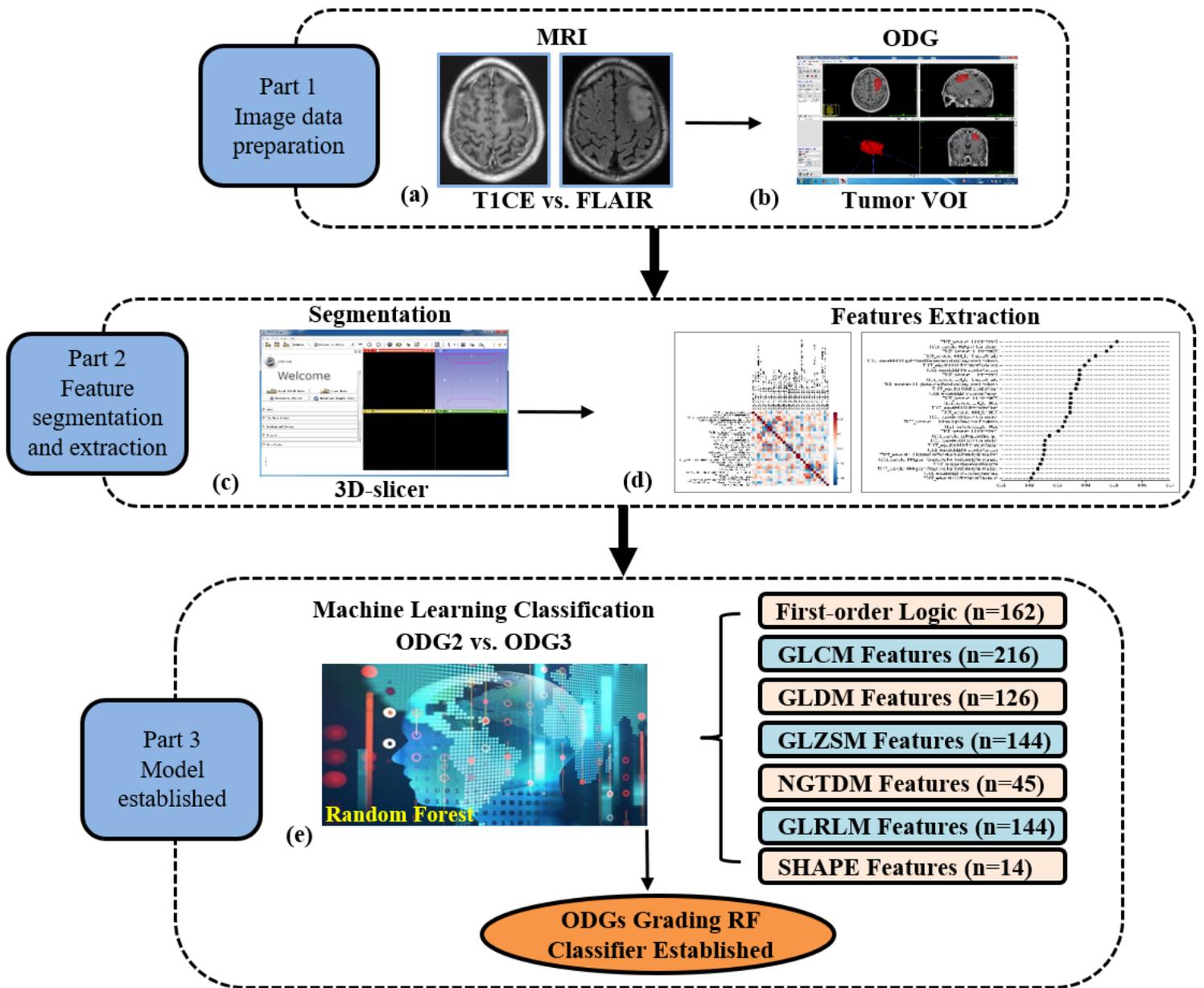


Figure 2

The main procedure of the radiomic strategy for preoperative ODGs grading. Based on T1CE and FLAIR data (a) and tumor volume of interest (VOI) manually drawn on resampled T1CE and FLAIR images (b), a group of parametric images are derived and the corresponding parametric maps of the whole tumor region are extracted (c). Utilizing radiomic features analysis; a big collection of tumor parameter attributes was acquired for the following machine learning process (d). Feature selection methods were implemented and compared using random forest (RF) classifier with additional discussion on model parameters to construct the optimal ODG grading model (e).

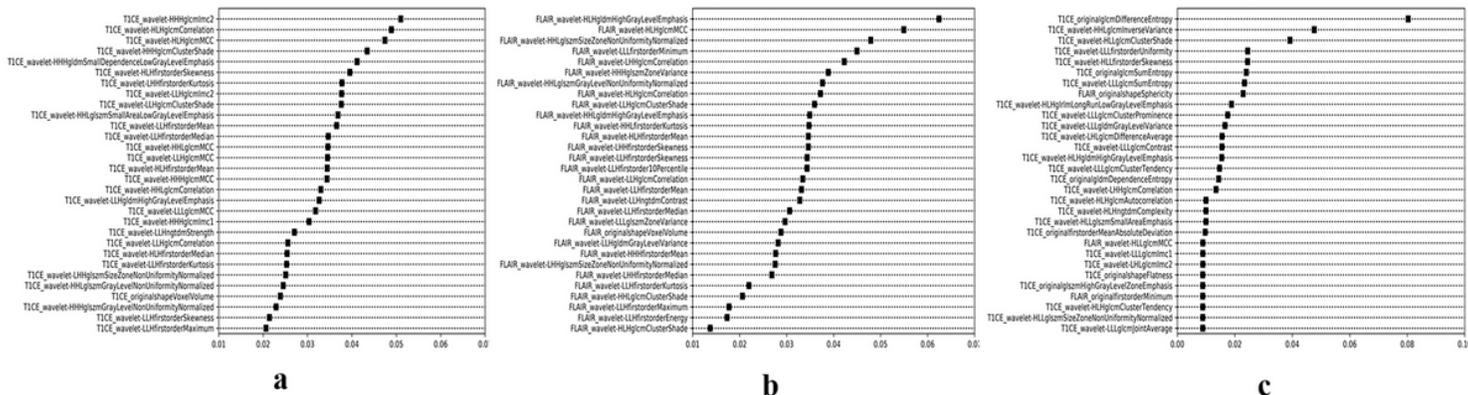


Figure 3

Feature importance plot shows mean decrease in Gini impurity. Features that most reduce Gini impurity are those that result in the least misclassification. Note: a = T1CE; b = FLAIR; c = T1CE+FLAIR.

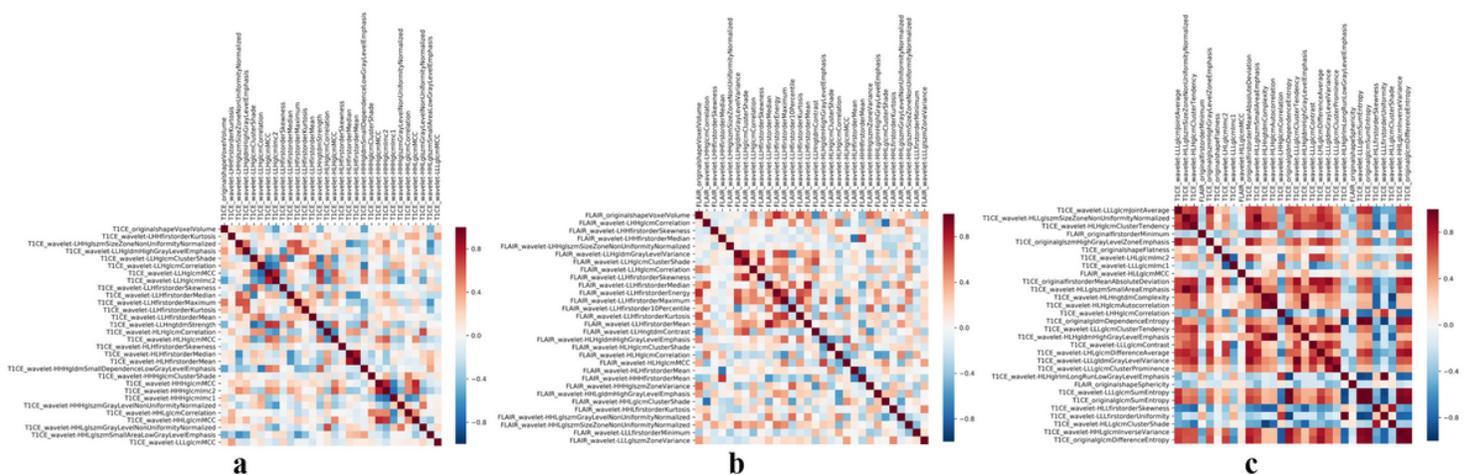


Figure 4

The radiomic heat map about the correlation analysis for feature selection: (a) T1CE; (b) FLAIR; (c) T1CE+FLAIR. Note: Red refers to positive correlations and blue refers to negative correlations. Different color depth indicates different values of correlation coefficients.

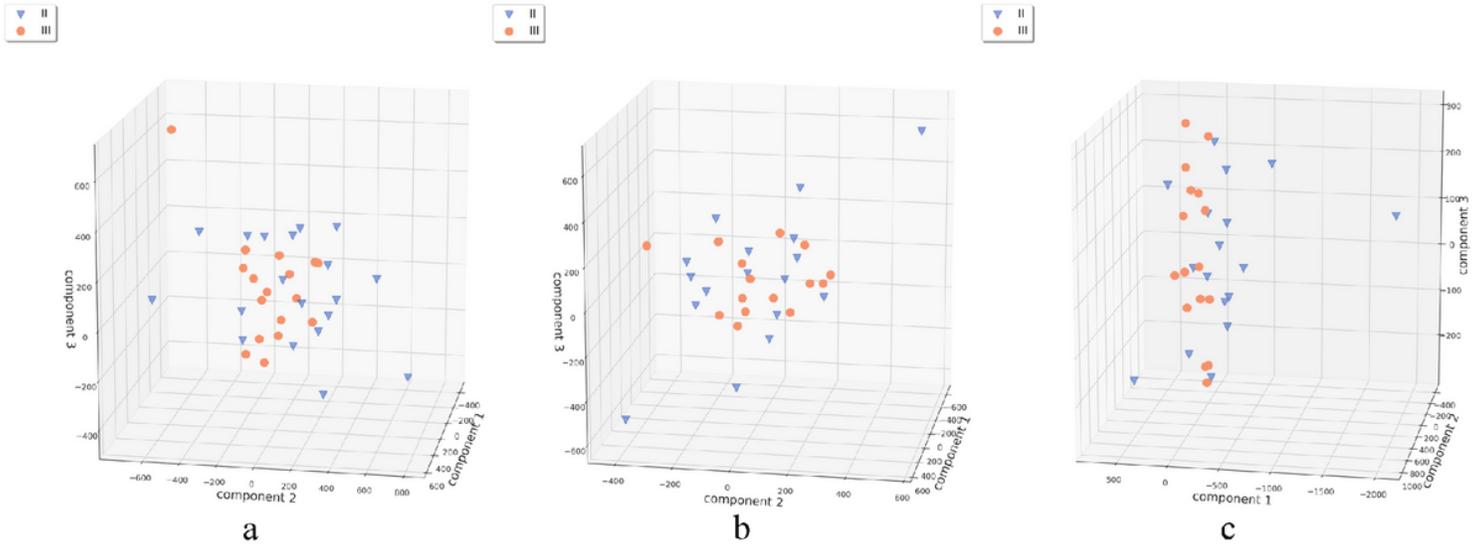


Figure 5

The calculated principal components for each tumor type were demonstrated based on the tumor tissue heterogeneity. II = ODG2, III = ODG3; component 1 = first principal component, component 2 = second principal component, component 3 = third principal component; a = T1CE; b = FLAIR; c = T1CE+FLAIR.

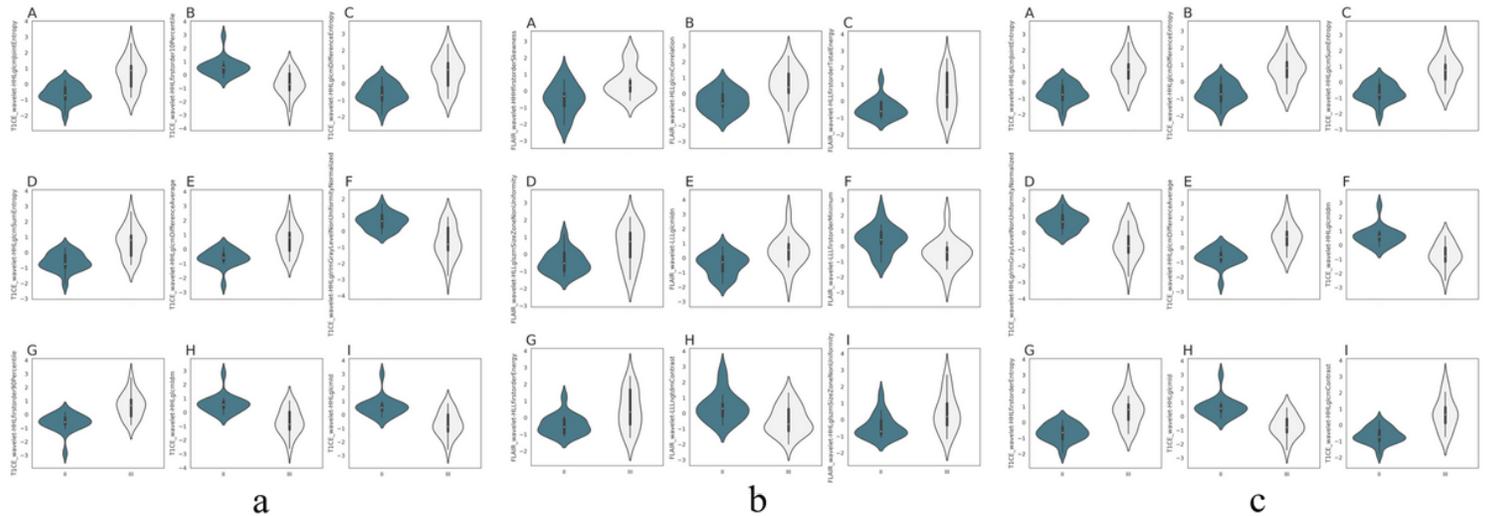


Figure 6

Violin plots show the values of first 9 radiomic features according to the grade of ODG. The small box in kernel density map represent the box plot. Points in small boxes = median values. Boundaries of small boxes = 25th and 75th percentiles. a = T1CE; b = FLAIR; c = T1CE+FLAIR. The violin represented kernel density map.

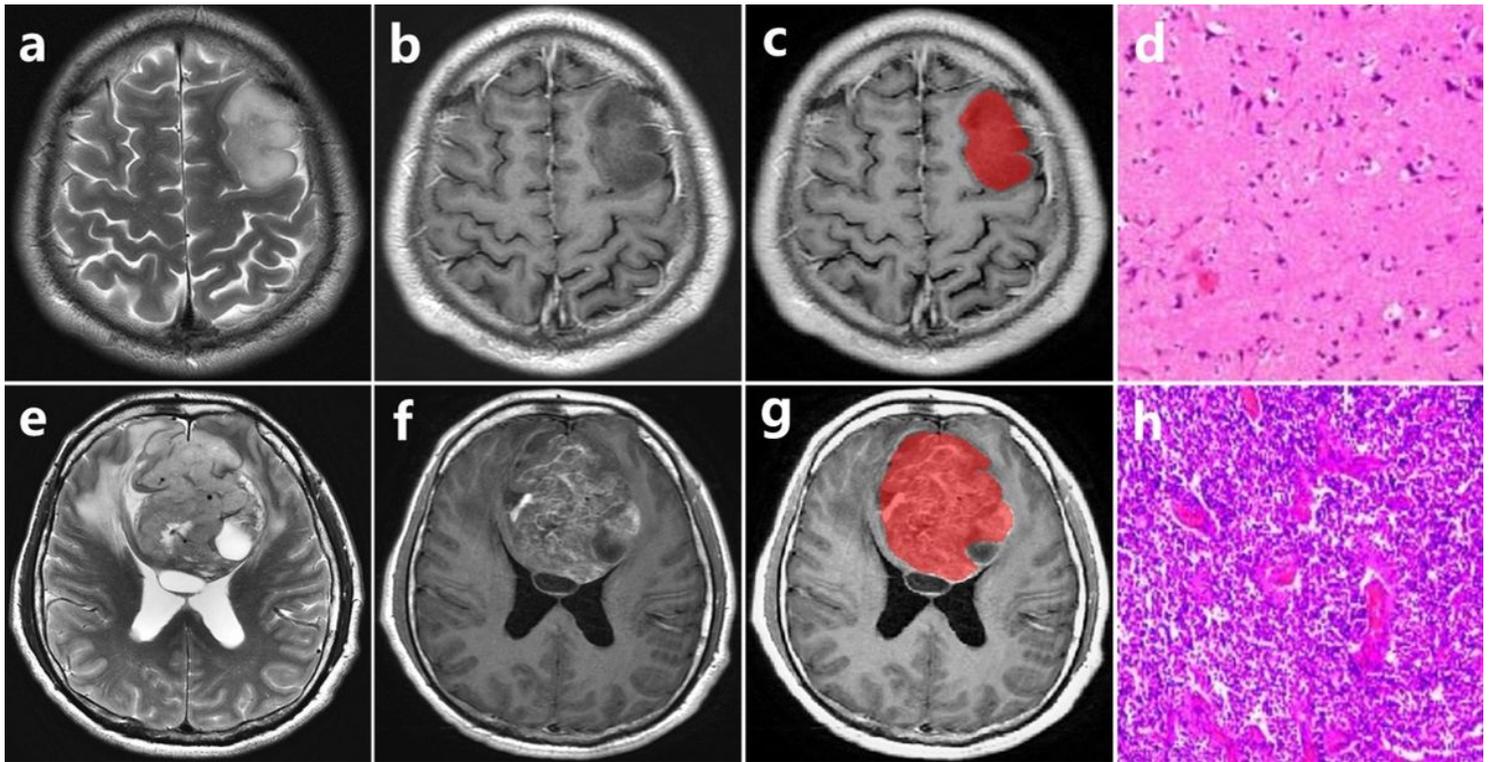


Figure 7

Upper row: ODG2 in the left frontal lobe from 33-year-old man; lower row: ODG3 in the bilateral frontal lobe from 46-year-old man. (a, e) T2-weighted image. (b, f) T1-weighted contrast-enhanced image. (c, g) The volume of interest of manually drawn. (d, h) Pathology slice images show cell density and vascular proliferation.