Comparable efficacies in differentiating WHO grade II from III oligodendrogliomas with machine-learning and radiologist’s reading

CURRENT STATUS: ACCEPTED

Sha-Sha Zhao
Xi'an Tangdu Hospital of No4 Military Medical University

Lin-Feng Yan
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

Xiao-Cheng Wei
GE Healthcare

Guang-Bin Cui
Xi'an Tangdu Hospital of No 4 Military Medical University

Wen Wang
Tangdu Hospital, Air Force Medical University

 wangwen@fmmu.edu.cn Corresponding Author
ORCiD: https://orcid.org/0000-0001-6473-4888

DOI: 10.21203/rs.2.9727/v1

SUBJECT AREAS
Neurology

KEYWORDS
Oligodendrogliomas; Machine learning; Radiomics; Random forest (RF); Magnetic resonance imaging (MRI)
Abstract
Abstract (236 words) Background: The medical imaging differentiation of World Health Organization (WHO) grade II (ODG2) from III (ODG3) oligodendrogliomas remains a challenge. We investigate whether combination of machine leaning with radiomics from conventional T1 contrast-enhanced (T1CE) magnetic resonance imaging (MRI) can offer improved efficacy. Methods: Thirty-six patients with histologically confirmed ODGs who underwent the T1CE MR examination before any intervention between January 2015 and July 2017 were recruited in the current study. The volume of interest (VOI) covering the whole tumor enhancement were manually drawn on the T1CE slice by slice using ITK-SNAP and a total of 1044 features were extracted from the VOI using Analysis-Kinetics software. Random forest (RF) algorithm and 5-fold cross validation were applied to differentiate ODG2 from ODG3. The diagnostic efficacies of radiomics-based machine learning and radiologist’s assessment were also compared. Results: Nineteen ODG2 and 17 ODG3 were included in the current study and ODG3 tended to present with prominent necrosis and nodular/ring-like enhancement (P < 0.05). The RF strategy with radiomics features produced the stable diagnostic efficiency, with an AUC, ACC, sensitivity, and specificity of 0.765, 0.763, 82.8% and 70.0%, 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 is comparable to that of radiologist. Conclusions: Machine-learning based on radiomics of T1CE offered comparable efficacy to that of radiologist on differentiating ODG2 from ODG3.

Background
Oligodendrogliomas (ODGs), predominantly occurring in adults with a peak between 40 and 60 years of age, constitute 5-20% of all gliomas together with mixed oligoastrocytoma [1], and patients with low-grade tumors (ODG2) being 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] is present in about 60–90% of ODGs, thus making it the molecular hallmark [1]. Calcification [4, 5] and a cortical-subcortical location [5, 6], most commonly in the frontal lobe [4], are regarded as the features of ODGs. In contrast to other low-grade gliomas, minimal to moderate enhancement and moderately increased perfusion is commonly seen in not only ODG3 but also ODG2,
making their differentiation complicated. Besides, ODG3 often has similar imaging features to ODG2 on conventional MRI, leading to no reliable tumor grade prediction. Edema, haemorrhage, cystic degeneration and contrast enhancement are more commonly seen in ODG3, reflecting histopathological findings, 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 and pathophysiological 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 low grade glioma (LGG), however, overlapped ADC values with that of ODG2 make DWI not a reliable maker to distinguish ODG2 from ODG3 [7]. Using a relative cerebral blood volume (rCBV) ratio cut-off value of 1.75, HGG can be differentiated from LGG with 95% sensitivity [8]. Unfortunately, these findings do not seem to apply to ODG and oligoastrocytoma, because markedly elevated rCBV can also be observed in ODG2, thus, a reliable distinction can not be consistently achieved [7, 9, 10]. This is due to the presence of the short capillary segments in ODGs [5] and may contribute to the relatively low specificity (70%) reported by Law et al[8]. (Focally) elevated rCBV therefore 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 MRI with machine learning algorithms, radiomics can provide comprehensive information that is difficult to perceive by visual inspection [12, 13] and is commonly used in the diagnosis, staging and prognosis of tumors [14-19]. However, most previous studies were largely focused on advanced MR techniques, the varied post-processing models, varied interpretation and evaluation criteria restricted their clinical applications. Besides their limited diagnostic powers, these advanced MRI techniques are not commonly available in some rural areas. However, the T1-weighted contrast-enhanced images (T1CE) is widely-used in almost all hospitals as a routine sequence for glioma diagnosis and staging. It is thus feasible to
combine radiomics with T1CE to establish a practical solution for differentiating ODG2 from ODG3.

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

Methods

Patients

This study was approved by our institutional review board and the requirement for informed consent was waived based on its retrospective nature. Between January 2015 and July 2017, consecutive patients with surgically confirmed ODG2 and ODG3 were retrospectively recruited. Tumors were classified according to 2007 WHO classification or 2016 WHO guidelines when enough information were available. The including criteria were, 1. patients underwent preoperative conventional MRI scan; 2. patients underwent gross total resection or subtotal resection of the lesion and a confirmative pathological diagnosis was made. Thirty-six patients 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). The patient selection process 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 conventional MR protocol included T1-weighted imaging (T1WI) performed before and after contrast enhancement, an axial T2-weighted imaging (T2WI), and a transverse fluid-attenuated inversion recovery (FLAIR). 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 excitation, 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; TI, 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 brain tumor imaging) re-examined the images and determined the final imaging diagnoses when inconsistency appeared between the two neuroradiologists. The preoperative conventional image features of tumor were retrieved based on the criteria outlined in **Supplementary Table 1** (online).

The VOI (volume of interest) 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 VOI covering the enhanced lesion were drawn slice by slice on T1CE, avoiding the regions of macroscopic necrosis, cyst, edema and non-tumor macrovessels [20].

**Radiomics Strategy**

**Feature extraction** Texture features include 42 histogram features, 432 gray level co-occurrence matrix (GLCM) features, 540 gray level run length matrix (GLRLM) features, 11 gray level size zone matrix (GLSZM) features, 9 form factor features and 10 Haralick features. A total of 1044 features were extracted from the T1CE images using Analysis-Kinetics (A.K., GE Healthcare) 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 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 (locally) 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 [21]. 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. Accuracy, sensitivity, specificity, positive
predicting value (PPV), and negative predicting value (NPV) 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 [22]. The whole
procedure of feature extraction and machine learning was depicted in **Figure 2**.

**Radiologist’s assessment** To compare the efficacies of neuroradiologist and machine learning in
differentiating ODG2 from ODG3, the images were also evaluated by three junior neuroradiologists
(X.L.F, G.X and Y.H with 6, 7 and 7 years of experience in neuroradiology, respectively). The
neuroradiologists were blinded to clinical information.

**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 image. 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 significantly frequent in ODG3 than in ODG2 groups (P = 0.044), reflecting the fact of 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. as a metric. Second-order texture measures of long run high grey level emphasis _All Direction_offset5_SD, correlation _All Direction_offset9_SD, and short run high grey level emphasis _All Direction_offset6_SD were among features that resulted in the largest decrease in the Gini index. Furthermore, patients with ODG2 had significantly higher long run high grey level emphasis _All Direction_offset5_SD, correlation _All Direction_offset9_SD, and short run high grey level emphasis _All Direction_offset6_SD were observed in ODG2 patients than ODG3 patients (Table 2). Finally, non-robust features such as the kurtosis computed from T1CE image and Gabor edge images, as well as the minimal intensity computed from T1CE image, were not among the top 10 relevant features.

The relative importance of the features computed by using the Gini index was shown in Table 2. Six
out of the 30 features were significantly different between ODG2 and ODG3 patients. 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 using all features.

**Diagnostic Performance of Radiomics and Radiologists**

The performance of radiomics and 3 radiologists in differentiating ODG2 from ODG3 was also compared (Table 3). The AUC of radiomics was 0.765 (95%CI 0.613-0.908), with a sensitivity of 0.828, specificity of 0.700 and accuracy of 0.763, 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 as good as the 3 junior radiologists. The representative cases of ODG2 and ODG3 were provided in Figure 5. The clinical application of radiomics-based machine learning could not be justified based on our findings.

**Discussion**

Digital radiological images are routinely acquired for almost every glioma patient, thus, medical imaging is rapidly becoming a crucial big data source for decoding tumoral phenotype. Radiomics has been suggested as a robust strategy to noninvasively classify lesions [14, 23]. This work suggests that radiomics from T1CE may be useful for differentiating ODG2 from ODG3, however, with the comparable efficacy to radiologists, thus, its clinical application could not be justified based on the current study.

From the angle of experiment design, there are three novel aspects 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. Finally, based on the social-economic consideration, the levels of accuracy were based on the radiomics of commonly available T1CE images, without an acquisition of spectroscopy, cerebral blood volume or perfusion information,
all of which would increase the scanning time and economic burden to patients. It is somehow out of our expectation that the radiomics strategy did not performed superior to radiologists. The reason for the not improved diagnostic performance of radiomics might be as the following. First, the information about ODG biology could not be fully reflected on the T1CE image. T1CE was based on the damage of brain blood barrier (BBB) as well as the proliferation of tumor blood vessels. However, the glioma invasion of normal tissue included the impaired white matter as well as mild alteration in blood perfusion or diffusion, which can be better revealed with advance MRI including DTI, DWI and ASL, etc. Second, T1CE offered some rough information to differentiate ODG2 from ODG3 can be visually inspected by radiologist. Because T1CE reflects the significant BBB damage and tumoral blood vessel proliferation, which are also quite obvious for the visual inspection. Not too much more detailed information than visual inspection was retrieved and contributed to the classifying, thus making the comparable efficacies of radiomics and radiologists. Even the radiomics strategy could not perform superior to radiologist, there is still potential to apply it as an adjunct to radiologist to overcome some problems attained to radiologists. First, 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 [24]. Second, fatigue adversely impacts the visual system including: worse accommodation, decreased saccadic velocity and reduced gaze volume and coverage [25]. Third, a number of cognitive biases may adversely affect the accuracy of a radiologists report of a glioma [26]. In order to reduce reporting time and cognitive biases, both of which may lead to reporting and diagnostic errors, radiomics offers a significant advantage [27], particularly in the context of general radiologists who may lack expertise in neuro-oncology. However, the current radiomics strategy involves too much pre- and post-process before the suitable machine learning model is established, thus, future studies focusing on the efficacy-cost balance of such a machine learning system should be conducted before its clinical application.

A few limitations of this work must be considered. First, the patient number is relative small and the robustness of the current strategy need to be consolidated on a future larger patient population. Second, the cross-sectional data used did not provide the ability to identify markers of survival [28].
future longitudinal study could provide such data. Third, the validation of the proposed strategy was
tested with 5-fold cross validation, instead of independent external dataset. A continuous effort on
enlarging the dataset so as to test its external validation is required.

Conclusions
In conclusion, this study demonstrates the use of a machine learning algorithm, derived from ‘real
word’ T1CE images, which can differentiated ODG2 from ODG3 in newly diagnosed gliomas with a
comparable efficacy to that of radiologists. The RF selected features in this study may reduce the
labor in applying this strategy, however, the clinical application of the proposed strategy can not be
justified based on our findings.

Abbreviations
ODG
HGG
LGG
rCBV
T1WI
T2WI
FLAIR
T1CE
RF
BBB

Oligodendroglioma
High Grade Glioma
Low Grade Glioma
Relative Cerebral Blood Volume
T1-weighted imaging
T2-weighted imaging
Fluid-Attenuated Inversion recovery
T1-weighted contrast-enhanced image
Random Forest
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 for publication
Our manuscript does not contain any individual persons data. (Not applicable)

Availability of data and material
Data sharing not applicable to this article as no datasets were generated or analysed during the
current study.

**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).

**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 and SZQ contributed to the data analysis and graph making, PD and CBY 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 |
| TQ  | MD | SYZ | MD   | ZJ  | MD |
| CGB | MD&Ph.D. | WW | MD&Ph.D. | GXW | MD |
|     |   |   |   | CSC | MD |
|     |   |   |   | WXC | MD |

**References**

1. Van Den Bent, M.J., J.E. Bromberg, and J. Buckner. Low-grade and anaplastic oligodendroglioma. Handb Clin Neurol. 2016; 134: p. 361-80.  
   http://doi.org/10.1016/B978-0-12-802997-8.00022-0.

2. Bromberg, J.E. and M.J. van den Bent. Oligodendrogliomas: molecular biology and
3. Jenkins, R.B., H. Blair, K.V. Ballman, 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): p. 9852-61.
http://doi.org/10.1158/0008-5472.CAN-06-1796.

4. Koeller, K.K. and E.J. Rushing. From the archives of the AFIP: Oligodendroglioma and its variants: radiologic-pathologic correlation. Radiographics. 2005; 25(6): p. 1669-88. http://doi.org/10.1148/rg.256055137.

5. Louis, D.N., H. Ohgaki, O.D. Wiestler, et al. The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol. 2007; 114(2): p. 97-109. http://doi.org/10.1007/s00401-007-0243-4.

6. Osborn, A.G., Osborn's Brain: Imaging, Pathology, and Anatomy (1st edition). 2012, Salt Lake City, UT: Amirsys, Inc.

7. Al-Okaili, R.N., J. Krejza, S. Wang, J.H. Woo, and E.R. Melhem. Advanced MR imaging techniques in the diagnosis of intraaxial brain tumors in adults. Radiographics. 2006; 26 Suppl 1: p. S173-89. http://doi.org/10.1148/rg.26si065513.

8. Law, M., S. Yang, H. Wang, 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): p. 1989-98.

9. Lev, M.H., Y. Ozsunar, J.W. Henson, 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): p. 214-21.
10. Chawla, S., S. Wang, R.L. Wolf, et al. Arterial spin-labeling and MR spectroscopy in the differentiation of gliomas. AJNR Am J Neuroradiol. 2007; 28(9): p. 1683-9. http://doi.org/10.3174/ajnr.A0673.

11. Lacerda, S. and M. Law. Magnetic resonance perfusion and permeability imaging in brain tumors. Neuroimaging Clin N Am. 2009; 19(4): p. 527-57. http://doi.org/10.1016/j.nic.2009.08.007.

12. Gillies, R.J., P.E. Kinahan, and H. Hricak. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016; 278(2): p. 563-77. http://doi.org/10.1148/radiol.2015151169.

13. Prasanna, P., J. Patel, S. Partovi, A. Madabhushi, and P. Tiwari. 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): p. 4188-4197. http://doi.org/10.1007/s00330-016-4637-3.

14. Huang, Y.Q., C.H. Liang, L. He, 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): p. 2157-64. http://doi.org/10.1200/JCO.2015.65.9128.

15. Horvat, N., H. Veeraraghavan, M. Khan, et al. MR Imaging of Rectal Cancer: Radiomics Analysis to Assess Treatment Response after Neoadjuvant Therapy. 2018; 287(3): p. 833-843. http://doi.org/10.1148/radiol.2018172300.

16. Tian, Q., L.F. Yan, and X. Zhang. Radiomics strategy for glioma grading using texture features from multiparametric MRI. http://doi.org/10.1002/jmri.26010.

17. Kalinli, A., F. Sarikoc, H. Akgun, and F. Ozturk. Performance comparison of machine learning methods for prognosis of hormone receptor status in breast cancer
tissue samples. Comput Methods Programs Biomed. 2013; **110**(3): p. 298-307. http://doi.org/10.1016/j.cmpb.2012.12.005.

18. Parmar, C., P. Grossmann, J. Bussink, P. Lambin, and H.J. Aerts. Machine Learning methods for Quantitative Radiomic Biomarkers. Sci Rep. 2015; **5**: p. 13087. http://doi.org/10.1038/srep13087.

19. Chae, H.D., C.M. Park, S.J. Park, S.M. Lee, K.G. Kim, and J.M. Goo. Computerized texture analysis of persistent part-solid ground-glass nodules: differentiation of preinvasive lesions from invasive pulmonary adenocarcinomas. Radiology. 2014; **273**(1): p. 285-93. http://doi.org/10.1148/radiol.14132187.

20. Yushkevich, P.A., G. Yang, and G. Gerig. ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. Conf Proc IEEE Eng Med Biol Soc. 2016; **2016**: p. 3342-3345. http://doi.org/10.1109/EMBC.2016.7591443.

21. Tagliamonte, S.A. and R.H. Baayen. Models, forests and trees of York English: Was/were variation as a case study for statistical practice. Language Variation & Change. 2012; **24**(2): p. 135-178.

22. Cui, Z., Z. Xia, M. Su, H. Shu, and G. Gong. Disrupted white matter connectivity underlying developmental dyslexia: A machine learning approach. Hum Brain Mapp. 2016; **37**(4): p. 1443-58. http://doi.org/10.1002/hbm.23112.

23. Aerts, H.J., E.R. Velazquez, R.T. Leijenaar, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014; **5**: p. 4006. http://doi.org/10.1038/ncomms5006.

24. Williams, L.H. and T. Drew. Distraction in diagnostic radiology: How is search through volumetric medical images affected by interruptions? Cogn Res Princ Implic. 2017; **2**(1): p. 12. http://doi.org/10.1186/s41235-017-0050-y.

25. Waite, S., S. Kolla, J. Jeudy, et al. Tired in the Reading Room: The Influence of
Fatigue in Radiology. J Am Coll Radiol. 2017; 14(2): p. 191-197.
http://doi.org/10.1016/j.jacr.2016.10.009.

26. Lee, C.S., P.G. Nagy, S.J. Weaver, and D.E. Newman-Toker. Cognitive and system factors contributing to diagnostic errors in radiology. AJR Am J Roentgenol. 2013; 201(3): p. 611-7. http://doi.org/10.2214/AJR.12.10375.

27. Thrall, J.H., X. Li, Q. Li, 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): p. 504-508. http://doi.org/10.1016/j.jacr.2017.12.026.

28. Jain, R., L.M. Poisson, D. Gutman, et al. Outcome prediction in patients with glioblastoma by using imaging, clinical, and genomic biomarkers: focus on the nonenhancing component of the tumor. Radiology. 2014; 272(2): p. 484-93. http://doi.org/10.1148/radiol.14131691.

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 ± SD                       | 45.6 ± 13.7 | 44.3 ± 15.1 | 45.0 ± 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)  |          |
| Feature                                    | ODG2                                      | ODG3                                      |
|-------------------------------------------|-------------------------------------------|-------------------------------------------|
| Gini Importance                           | 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 |
| 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) |
| 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) |
| Short Run High Grey Level Emphasis_All Direction_offset6_SD | 0.94   | 23.26 | 17.65-29.72 | 14.42 | 5.94-21.34 |
| RMS                                       | 0.92   | 930.73 | 818.53-1037.73 | 1213.40 | 949.17-1493.34 |
| Long Run High Grey Level Emphasis_All Direction_offset7_SD | 0.90   | 469.55 | 285.36-601.49 | 285.89 | 178.74-362.53 |
| 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) |
| 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) |
| Feature                          | Mean   | Min   | Max   | SD     | Mean   | Min   | Max   | SD     |
|---------------------------------|--------|-------|-------|--------|--------|-------|-------|--------|
| GLCM Entropy_All Direction_offset5_SD | 0.58   | 0.32  | 0.09-0.38 | 0.18 | 0.11-0.20 |
| Compactness2                    | 0.52   | 16.16 | 15.23-17.10 | 19.77 | 15.86-26.80 |
| Haralick Correlation_angle135_offset9 | 0.52   | 5470.3 (10⁵) | 2510.0-7750.0 (10⁵) | 4880.59 (10⁵) | 2305.0-4335.0 (10⁵) |
| Run Length Nonuniformity_All Direction_offset3_SD | 0.49 | 125.50 | 42.61-117.46 | 180.04 | 20.69-327.44 |
| Intensity Variability            | 0.45   | 11.45 | 7.98-14.75 | 9.04 | 2.75-11.69 |
| 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) |
| Age                             | 0.44   | 46    | 43-55 | 44    | 37-57 |
| 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) |
| High Grey Level Run Emphasis     | 0.42   | 4.26  | 2.64-5.14 | 6.77 | 3.02-9.50 |
| Short Run Emphasis_All Direction_offset8_SD | 0.36 | 5.42  | 2.48-7.41 | 7.21 | 2.74-11.59 |
| GLCM Entropy_All Direction_offset7_SD | 0.39 | 7.69  | 6.66-8.70 | 8.18 | 6.27-10.19 |
| GLCM Entropy_angle135_offset9    | 0.39   | 130.87 (10⁻⁵) | 63.37-175.07 (10⁻⁵) | 110.70 (10⁻⁵) | 66.12-147.02 (10⁻⁵) |
| Minimum Intensity               | 0.38   | 114.79 | 17.0-145.0 | 149.82 | 7.0-275.0 |
| 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) |
| High Grey Level Run Emphasis     | 0.36   | 5.42  | 2.48-7.41 | 7.21 | 2.74-11.59 |
| 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) |
| GLCM Entropy_angle90_offset8     | 0.34   | 8.27  | 7.19-9.14 | 8.11 | 6.46-9.73 |
| Cluster_Shade_angle135_offset3   | 0.29   | 28109.70 | -47175.30-65223.50 | 250959.5 | -23039.0-229078.0 |
| Cluster_Shade_All Direction_offset3_SD | 0.29 | 27784.98 | 4378.85-26851.50 | 29946.8 | 6717.39-22431.35 |
| GLCM Entropy_All Direction_offset9_SD | 0.26 | 0.31  | 0.18-0.37 | 0.22 | 0.13-0.35 |

**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.
|         | Sensitivity | Specificity | AUC                      | ACC  |
|---------|-------------|-------------|--------------------------|------|
| Radiomics | 0.828       | 0.700       | 0.765 (95% CI: 0.613, 0.908) | 0.763 |
| 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**

![Flow diagram of the study design.](image-url)

**Figure 1**

Flow diagram of the study design.
The main procedure of the radiomics strategy for preoperative ODGs grading. Based on T1CE data (a) and tumor volume of interest (VOI) manually drawn on resampled T1CE image (b), a group of parametric images were derived and the corresponding parametric maps of the whole tumor region were extracted (c). Utilizing histogram features, GLCM features, RLM features, GLSZM features, Form Factor features and Haralick 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 showed mean decrease in Gini impurity. Features that most reduce Gini impurity were those that result in the least misclassification.
Figure 4

The radiomic heat map about the correlation analysis for feature selection. Note: Blue refers to positive correlations and red refers to negative correlations. Different color depth indicates different values of correlation coefficients.
Figure 5

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 showed cell density and vascular proliferation.